Fall 2012

Identifying and Characterizing Non-Coding RNAs in the Dinoflagellate *Karenia brevis*

Helen Namataka
University of Southern Mississippi

The University of Southern Mississippi

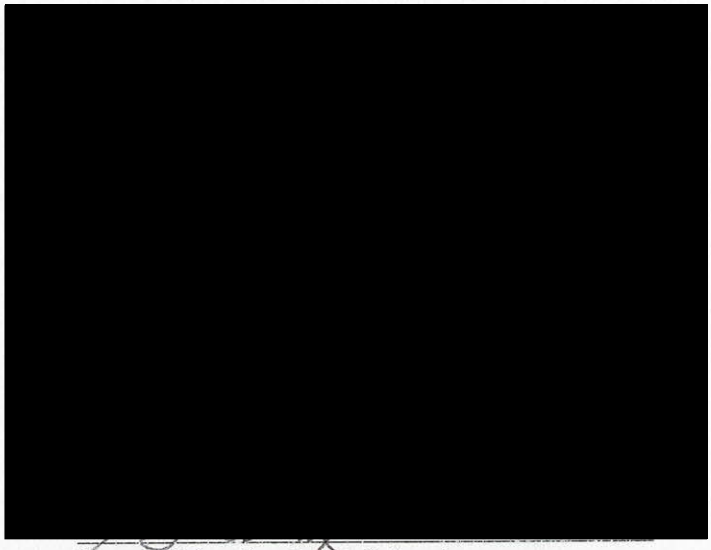
IDENTIFYING AND CHARACTERIZING NON-CODING RNAs IN THE DINOFLAGELLATE *KARENIA BREVIS*

by

Helen Namataka

A Thesis
Submitted to the Graduate School
of The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Master of Science

Approved:



Dean of the Graduate School

December 2012

ABSTRACT

IDENTIFYING AND CHARACTERIZING NON-CODING RNAS IN THE DINOFLAGELLATE *KARENIA BREVIS*

by Helen Namataka

December 2012

When algal cells proliferate and accumulate in marine and fresh water systems, they form algal blooms. The majority of these blooms are beneficial, but a significant number are detrimental and are known as harmful algal blooms (HABs). A number of negative effects, including closing of recreational beaches and economic loss, are observed during HABs. Predictably, the longer a bloom persists, the greater its effects on human, environmental and economic health. *Karenia brevis*, a mixotrophic dinoflagellate, forms HABs, and blooms caused by this organism have been known to remain several months after formation. For these reasons, research has been conducted to discover those factors that are responsible for the formation, maintenance and terminations of *K.brevis* blooms. These factors, whether man-made (e.g. eutrophication) or naturally occurring (e.g. aquatic fronts), have been and continue to be extensively investigated. As part of an effort to find molecular or genetic determinants that control *K. brevis* biology, it was discovered that these cells express a number of anti-sense RNAs. For the first time in *K.brevis*, I have confirmed the expression of targeted anti-sense RNAs (asRNAs) and characterized their sequences.

We have developed a model in which the expression of asRNAs leads to the formation of double-stranded RNAs (dsRNA). I have immunofluorescence data to support the presence of dsRNA in *K. brevis* cells. We hypothesize, based on the known functions of asRNAs in other systems, that the presence of asRNAs (and hence dsRNAs) helps to regulate gene expression at a post-transcriptional level.

ACKNOWLEDGMENTS

I would like to express my utmost gratitude to my advisor, Dr. Timothy McLean whose indispensable expertise, understanding, and enormous patience were instrumental in my graduate experience. This thesis would not have been accomplished without his support and shared skill. I would like to thank my committee, Dr. Mohamed Elasri and Dr. Glen Shearer, whose patient instruction and knowledge were invaluable.

For their selfless support, a special thank you goes to some remarkable individuals in the Department of Biological Sciences; David Jayroe, Maria Basco, Gyan Sundar Shukha, Chris Flood, Rebecca Browning, Scott Anglin, Davida Crossley, and to Baobin Wang for her assistance with my immunofluorescence project.

Finally, I would have been lost without tireless encouragement from my friends and family, particuarly my parents, Basil and Betty Wanzala, whose prayers and unwavering devotion to my success have carried me to this point and especially through this process.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF ILLUSTRATIONS

Figure

CHAPTER I

INTRODUCTION

The mixotrophic dinoflagellate *Karenia brevis* is recognized for its role in the formation of harmful algal blooms (HABs) along all Gulf of Mexico (GoMex) coastlines (NCCOS, 2006). *Karenia brevis* is a eukaryote that was classified under the five-kingdom system as belonging to the kingdom Protista, but in a 2005 reclassification, was placed under the kingdom Chromalveolata (Guiry & Guiry, 2012). Formerly known as Gymnodinium breve and Ptychodiscus brevis, in November 2000 this organism was renamed by Danish scientists Gert Hansen and Ojvind Moestrup in honor of Karen Steidinger, the American biologist whose revolutionary work on *Karenia brevis* has greatly facilitated its demystification (FWRI, 2009; Lindstrom, 2007). One key reason for the recent classification of Gymnodinium breve as *Karenia brevis* can be explained by the morphology of this organism. Unlike other members of the genus Gymnodinium, *Karenia brevis* lacks a cellular covering (it is described as athecate) that renders it especially fragile (FWRI). This feature, which allows the cells to be broken open easily, is one of the key contributory phenomenom leading to the intoxicating effects associated with this dinoflagellate (Derby, Galliano, Krzanowski, & Martin, 2002; Hackett, Anderson, Erdner, & Bhattacharya, 2004). Once K.brevis ruptures, it releases neurotoxins, known as brevetoxins, that are harmful to fish and marine mammals. Other effects of these toxins will be elucidated in detail elsewhere in this manuscript.

With a straight groove at the anterior section of the cell that extends to its ventral and dorsal surfaces, the structure of *Karenia brevis* (Figure 1) has been likened to a cloverleaf (FWRI, 2009; Lindstrom, 2007). The cell is approximately 18-

45 μm in size, compressed dorso-ventrally and concave shaped to its ventral side (FWRI, 2009). In comparison, red blood cells measure about 8μm, and human hair follicles range in size from 16-35μm. The motility of this unicellular marine organism is achieved with the use of a longitudinal flagellum for propulsion and a rotary or transverse flagellum that allows for changes in direction. Cells display diel vertical migration by swimming near the surface during the day to photosynthesize and migrating to deeper levels at night for dissolved nutrients (McKay, Kamykowski, Milligan, Schaeffer, & Sinclair, 2006).

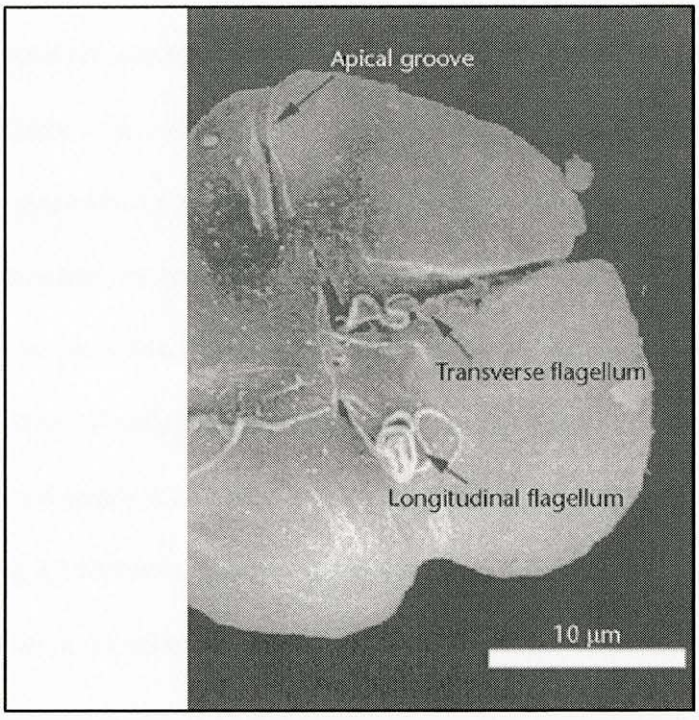


Figure 1. Structure of a Karenia brevis cell. Photo credit: Heimann, K. (2012). Photo retrieved, June 3, 2012 from: http://www.els.net/Wiley. Electron micrograph image shows cloverleaf shape of K. brevis cell with apical groove and transverse and longitudinal flagella.

Dinoflagellates have a fascinating evolutionary history. For example, whereas other photosynthetic organisms contain plastids that originated via primary or secondary endosymbiosis, dinoflagellates acquired their plastids by tertiary endosymbiosis (Yoon et al., 2005). In primary endosymbiosis, eukaryotic cells engulfed a cyanobacterium in their cytoplasm and the engulfed photosynthetic prokaryote eventually evolved into a double membrane bound plastid. On the other hand, secondary endosymbiosis was a result of non-photosynthetic protists engulfing eukaryotic algae forming plastids that are bound by 3-4 membranes. Finally, tertiary endosymbiosis by dinoflagellate plastids was achieved by engulfing algae containing a secondary plastid (Nosenko et al., 2006; Yoon et al., 2005). While a majority of dinoflagellates have a plastid containing the common accessory pigment peridinin, Karenia spp. acquired fucoxanthin-containing plastids by engulfing haptophyte algae instead. These plastids are bound by 3 membranes containing two fucoxanthin derivatives and no peridinin (Hackett et al., 2004; Yoon, Hackett, & Bhattachrya, 2002). Additionally, dinoflagellates are involved in horizontal gene transfer where nuclear and plastid genes from other organisms are transferred to the dinoflagellate nucleus, causing a transformation of the genome (Nosenko & Bhattacharya, 2007). In peridinin-containing dinoflagellates, apart from a few light dependent reaction genes that encode important subunits of the photosystem, e.g. cytochrome bf and ATP synthase complexes, and still remain in the plastid, almost all other organellar genes have been transferred to the nuclear genome. In addition, whereas plant chloroplast genomes have their genes linked on a single large circular genome containing up to 250 genes, peridinin-containing dinoflagellate chloroplast genes have been found

existing on small circular molecules or minicircles, each one having only one or two genes (Howe, Nisbet, & Barbrook, 2008; Zhang, Cavalier-Smith, & Green, 2002). Such massive gene transfer and plastid genome rearrangement has not been intensively investigated in fucoxanthin-containing dinoflagellates, such as K. brevis. In Karenia brevis, it has been shown that the nuclear-encoded oxygen-evolving enhancer 1 protein (Psb0) and the plastid-targeted GAPDH were replaced by their homologs from the engulfed haptophyte (Archibald, 2005; Hackett et al., 2004; Yoon et al., 2005). Nosenko et al. (2006) postulated that the ability to recruit new genetic material allows chromalveolates to adapt quickly to environmental changes. For instance, unlike diatoms, Karenia brevis is incredibly tolerant to low iron conditions (Nosenko et al., 2006). Results obtained by Nosenko et al. (2006) indicate that K. brevis contains the iron-sulfur protein ferredoxin and acquired the non-iron protein flavodoxin from red algae. Since existing evidence shows that as a mechanism of adaptation, many algae display a physiological switch from ferredoxin to flavodoxin under low iron conditions, Nosenko et al. (2006) surmised that having endosymbiotically acquired the non-iron protein, K. brevis cells are able to survive under conditions of iron limitation. Ultimately, these processes provide clues into how the eukaryotic, and particularly K.brevis, genome evolved and advance our understanding of how K. brevis adapts to both internal and external environmental changes (Nosenko & Bhattacharya, 2007; Yoon et al., 2005).

Perhaps the most intriguing dinoflagellate features are those related to their immense genomes. Whereas other eukaryotic algae contain an average of 0.54 pg DNA/cell, dinoflagellate genomes contain 3-250 pg DNA/cell (Hackett et al., 2004).

With roughly 1 x10^{11} base pairs (~100 pg/cell), the K. brevis genome is about 30 times that of humans whose genome contains 3.4 x10^{9} base pairs or 3 pg/cell (Hackett et al., 2004; Lidie, Ryan, Barbier, & Van Dolah, 2005; Van Dolah et al., 2008). Karenia brevis cells are haploid with 121 chromosomes. Dinoflagellate chromosomes exist in a liquid crystal state, and they are permanently condensed throughout the cell cycle except for rare loops of DNA that extend outside of the condensed chromosome core and are thought to be sites of active transcription (Morey et al., 2011; Van Dolah et al., 2008; Walker, 1982). Karenia brevis lacks nucleosomes and like all other dinoflagellates, K. brevis was thought to lack histones as well, using histone-like proteins for DNA packaging instead (Hackett et al., 2004). However, recent data indicates that all 4 core histone proteins are present in some dinoflagellates (e.g. Alexandrium catenella, Alexandrium tamarense, and Crypthecodinium cohnii) and that the difficulty in detecting these protein in other dinoflagellates such as K. brevis may be due to low expression levels as was the case for A. catenella (Hackett et al., 2005; Lin et al., 2011). Dinoflagellates are also unique in that, unlike other eukaryotes, their DNA contains a fifth base, 5-hydroxymethyluracil, a modified nucleotide that replaces 12-70% of thymines, depending upon the species of dinoflagellate (Hackett et al., 2004, 2005; Lin, 2011; Van Dolah et al., 2008). Curiously, typical eukaryotic transcriptional elements such as the TATA box have not yet been identified in dinoflagellates, suggesting an alternative mechanism for regulating transcription and gene expression. Findings such as a TBP-like protein in C. cohnii and a K. brevis EST with some similarity to the stabilizing general factor, TFIIE, have only lead to more questions concerning gene

regulation in dinoflagellates (Lidie et al., 2005; Lin, 2011). Nonetheless, expression studies from the Van Dolah lab (NOAA-Charleston, SC) revealed that genes involved in the translational machinery far out numbered those involved in transcription, implying a move to post-transcriptional regulation in K. brevis (Lidie et al., 2005). Identification of a spliced leader sequence in all the dinoflagellate species (including Karenia brevis) lends support to the notion that dinoflagellate genes are regulated post-transcriptionally. In K. brevis, the dinoflagellate spliced leader was found to be a 22 nucleotide sequence ligated onto the 5' end of nuclear mRNAs from a putative spliced leader RNA containing the splice donor site immediately adjacent to the leader sequence (Lidie & Van Dolah, 2007a; Lin, 2011). In other dinoflagellates such as Pfiesteria piscicida, trans-splicing converts polycistronic pre-mRNAs or tandem-repeated transcripts into mature monocistronic mRNAs. Alternatively, it has been proposed that in K. brevis, trans-splicing serves to regulate gene expression post-transcriptionally by ensuring that only transcripts that have the spliced leader sequence are translated (Lidie & Van Dolah, 2007a; Lin, 2011; Zhang et al., 2007). Moreover, Lin (2011) observed that low histone expression levels coincide with the presence of rare transcripts for the transcriptional mechanism, which supports the above model for gene regulation,since histones play a critical role in transcriptional regulation, at least in other organims. This creates a potential for translational and post-translational modes of gene regulation including protein phosphorylation, methylation and sumoylation, all of which remain poorly studied in dinoflagellates (Lin, 2011).

CHAPTER II

REVIEW OF RELATED LITERATURE

Formation and Maintenance of Harmful Algal Blooms

HABs arise when micro- or macroalgae accumulate to high levels and have a damaging effect on the environment and its inhabitants. As a direct result of the red discoloration of water that is frequently associated with the accumulation of planktonic species such as cyanobacteria, diatoms and dinoflagellates, these HABs are commonly referred to as red tides (Matthews & Brandt, 2004). One early description for the result achieved when dinoflagellates reach such high concentrations as to effect a change in color defines the term red tide as "the sudden proliferation of microorganisms, typically in the marine environment, that give rise to discolored water and other phenomena" (Martin & Martin, 1976). The varying colors of a HAB can be explained by the different pigments of the photosynthetic species involved in the formation of the bloom (Matthews & Brandt, 2004). In fact, the color of a typical K. brevis bloom ranges from brown, burgundy, red, and even yellow as a result of the pigmentation attributed to fucoxanthin and depending on such factors as density of the bloom, accompanying dissolved and/or suspended materials in the water, and light intensity. While some harmful algae do not produce toxins, blooms formed by these non-toxic species can be quite detrimental by diminishing or obstructing the depth of light penetration which subsequently reduces aquatic vegetation essential to the growth of certain fish and shellfish (FWRI, 2009). Additionally, as blooms, toxic or non-toxic, decline there is a marked decrease in dissolved oxygen levels due to decomposition of algae by bacteria. Evidence for this

oxygen depletion is shown in mortalities of both fish and aquatic plant life (Rabalais, 2004). Research from past and recent *K. brevis* HAB events suggests that anoxic conditions cause almost as much damage to marine life as brevetoxin intoxication (USEPA, 2008). In 1971, mortalities of fish, shellfish, porpoises, barnacles and turtles as well as destruction of entire benthic populations were attributed to reduced dissolved oxygen levels after a red tide event in Tampa Bay, Florida (Martin & Martin, 1976). Likewise, following the decline of a particularly devastating K. brevis bloom that had left many Florida manatees dead in the summer of 2005, many more animals were killed as a result of dissolved oxygen deficiency in waters covering over 2,162 square miles (USEPA, 2008).

Until the 1970s, harmful algal blooms were restricted to European, Japanese and North American shorelines. By the early '90s, increasing HAB events had been detected along the coasts of Kenya, South Africa, India, the Philippines, New Zealand and Australia (Hallegraeff, Valentine, Marshall, & Bolch, 2004). Since then, blooms have been reported worldwide. Improved detection techniques, increased scientific interest, an explosion in aquaculture with its associated pollution, and transport via ship ballast waters are some hypotheses that have been cited for this increase in HAB proliferation (Anderson, 1989). A case in point, HABs caused by the pelagophyte, Aureococcus anophagefferens, were recorded for the first time along the mid-Atlantic coast of the United States in 1985 where the high density of the bloom hindered the ability of shellfish to filter feed and led to the shellfish starving to death (Anderson, 1989). A similar occurrence was later observed at Saldanha Bay in South Africa where mariculture activities have suffered as a result of ongoing Aureococcus blooms

since the species was first recorded on the South African Bay in 1988. It is thought that A. anophagefferens was transported to South African coastal waters from the United States through ballast waters (Matthews & Brandt, 2004). The National Oceanic and Atmospheric Administration (NOAA) reports that although little is understood about the causes of harmful algal blooms, evidence from research in this area has been found linking population explosions of algal species to various environmental and human factors (Hallegraeff et al., 2004). Some of the well researched growth conditions that determine selective HAB proliferation in one environment over another have included differences in salinity, temperature and light. Traditionally, K. brevis bloom development is supported by high salinity levels of >24 psu. However, in a study carried out to determine the cause of a 1996 algal bloom along the Northern Gulf of Mexico (NGoMex) consisting of Mississippi, Alabama, and Louisiana coastlines, it was discovered that salinities along the cited regions were well below 24 psu (Brown et al., 2005). The phenomenon was determined to be a result of a combination of factors including tropical storm Josephine that facilitated dinoflagellate transport to and along the NGoMex shores (Brown et al., 2005). Since 1996, no algal blooms have been reported along these coastal areas, and HAB events due to K.brevis continue to be seen in areas with salinities >24 psu. Researchers concluded that although K.brevis clearly prefers higher salinity, a multitude of external factors are capable of allowing existing blooms to persist in lower salinity waters and as a result, bloom monitoring should be maintained even in low risk areas (Brown et al., 2005).

Increased nutrient runoff continues to be cited as a sustaining factor for the formation and/or maintenance of HABs. K. brevis blooms occur most frequently along the West Florida Shelf (WFS), one of the widest continental shelves in North America. This continental plate extends about 200 km off the coast of Florida where variations in the loop current affect the water. On the other hand, in the innermost shelf where HABs are most prevalent, waters are influenced by wind and land runoff (Brand & Compton, 2006). Using spatial and temporal pattern comparisons in the frequency of K.brevis blooms from 1954-1963 and 1994-2002, Brand and Compton (2006) hypothesized that an accumulation of nutrients from agricultural runoff is essential to bloom formation. They offer several theories for increased nutrient runoff, some of which include a rise in human occupation along the WFS, release of buried nutrients and reduced advection. Evidence shows that the occurrence of K.brevis blooms along the WFS is directly proportional to the persistent increase of human activities in this area (Brand & Compton, 2006). A case to support this occurrence is the drainage of the northern Everglades from the 1880s to the 1960s when the waters that maintained this wilderness were instead rerouted to the Atlantic Ocean and in later years to the GoMex (McPherson, 2003). Deprived of a constant supply of water and nutrients, the newly drained area was quickly desiccated. However, with the use of fertilizers, a significant portion of the cleared lands was effectively turned into agricultural territory (McPherson, 2003). The combined effort had the two-fold effect of introducing nutrients that would eventually flow into the GoMex as well as releasing nutrients that had until that time been buried in the everglades. Other factors such as reduced advection that have enhanced

eutrophication are caused by barriers to water transport leading to nutrient build-up (Yang & Weisberg, 1999). In places where K.brevis blooms are predominant, accumulation of nutrients over time continues to support bloom formation and maintenance (Brand & Compton, 2006).

Although previously thought to be a strictly autotrophic species, Karenia brevis has only recently been shown to be mixotrophic (Jeong et al., 2005). The cells rely on photosynthesis as a major source of nutrition but they are also important grazers. Like other alveolates, these dinoflagellates are able to heterotrophically acquire energy by feeding on picophytoplankton. Moreover, it is entirely possible that fucoxanthin dinoflagellates like K. brevis acquired their tertiary endosymbiont by a heterotrophic mode of feeding (Jeong et al., 2010). Jeong et al. (2005) successfully showed for the first time that K. brevis is an important predator of the cyanobacterium Synechococcus sp. and proposed a link between grazing of K. brevis on Synechecoccus and bloom maintenance. K. brevis blooms have been known to form in offshore waters where anthropogenic nutrient concentrations are low but Synechecoccus biomass is high. Jeong et al. (2005) suggest that perhaps one reason for persistence of K. brevis blooms is the potential to feed on picophytoplankton species such as Synechecoccus even under conditions of limited inorganic nutrients. Experiments carried out by Seong, Jeong, Kim, Kim, & Kang (2006) indicated that certain red-tide algae feed on heterotrophic bacteria which have a high phosphorous content. Furthermore, recent laboratory studies showed that grazing on Synechecoccus sp. contributed up to 40% of the inorganic nitrogen required per hour for growth of K. brevis cells through nitrogen fixation (Gilbert et al., 2009). If these

results can be verified in the field then the cyanobacteria would provide an additional source for both dissolved nitrogen and phosphorous, important nutrients for K. brevis bloom formation and maintenance, especially in offshore waters where these nutrients are limited. Finally, this capacity for heterotrophy suggests that K. brevis is able to grow and sustain blooms even under limited light conditions as is often the case when the majority of cells would be self-shaded in dense blooms (Gilbert et al., 2009).

Brevetoxins

Karenia brevis ruptures to release a suite of fused polyether ladder molecules known as brevetoxins that can act as neurotoxins in vertebrates. Once released, the potent toxins not only cause disruption of aquatic ecosystems due to the major role brevetoxin poisoning plays in untold fish and marine mammal mortalities but the neurotoxins are hazardous to human health as well (Lindstrom, 2007). As many as nine brevetoxin congeners (PbTx-1 to PbTx-9) have been documented and all active forms are related in that they are based on the two parent backbones Brevetoxin A and Brevetoxin B (Baden, Bourdelais, Jacocks, Michelliza, & Naar, 2005). Brevetoxins function by binding to the α-subunit of voltage sensitive sodium channels (VSSCs) in neuronal cells and persistently activating them as a consequence. VSSCs are integral membrane proteins that function to conduct sodium ions across a cell's plasma membrane in response to changes in voltage. Similarly, the α-subunit is a polypeptide glycoprotein on the VSSC that is composed of four homologous domains, each of which contains six trans-membrane α-helices (S1-S6). For each domain, three of these helices are neutral in charge, two are hydrophobic and one, S4, carries a highly positive charge. Because this S4 α- helix is extremely sensitive to

altering membrane voltage, it undergoes an allosteric change upon slight shifts in transmembrane potential. When all four S4 helices exhibit allosteric realignment, they cause a conformational change in the VSSC that causes the channel to open (Baden et al., 2005). Voltage dependant channels are open upon activation allowing an influx of sodium ions and are closed when deactivated to stop the ion influx. Specialized cells such as neurons and certain muscle cells produce and conduct electric impulses that cause voltage changes across the plasma membrane (Lodish et al., 2004). This voltage change, known as an action potential, is characterized by a depolarization stage in which the membrane potential is approximately +50mV due to the inside of the cell being more positively charged than the exterior as sodium ions pass into the cystol. Depolarization is immediately followed by a repolarization phase in which the cell is at rest with a resting potential of about -60mV and the cell's interior is negatively charged relative to the outside. Depolarization causes a conformational change in the voltage-gated sodium channels that allows the channels to open whereas they remain closed during repolarization (Lodish et al., 2004). By attaching to voltage-gated sodium channels, brevetoxins inhibit the channels' normal response to repolarisation and depolarization. Instead, the toxins cause a conformational change that favors negative active potentials and allows uncontrolled and enhanced influx of sodium ions into the cell and, as a consequence, persistent activation of neuronal and skeletal muscle cells (Baden et al., 2005; NIEHS, 1996; FWRI, 2009). Human intoxication through consumption of brevetoxin-contaminated shellfish is therefore characterized by neuronal and muscular symptoms such as reversal of sensation to hot and cold temperatures, muscle pain and tingling.

Gastrointestinal symptoms such as abdominal pain, vomiting, diarrhea and nausea are also common. Brevetoxins are tasteless and odourless, characteristics that severely impair their detection, particularly in the absence of a bloom. In addition, these lipid soluble nerve toxins are heat and acid labile, meaning that even proper food preparation is hopelessly ineffective as a measure against brevetoxin poisoning (Baden, Rein, Gawley, Jeqlitsch & Adams, 1994).

Aerosolized brevetoxins have been implicated in cases of respiratory irritation that may have particularly severe health effects in individuals with prior respiratory problems such as asthma (Fleming et al., 2007). Woodcock (1948) recorded remarkable disparities between droplets of vapour collected from normal sea water and during a red tide event. The aerosolized droplets from the red tide contained greenish particles analogous in structure to Karenia brevis. This provides evidence that aerosols from a red tide event carry both the brevetoxins and the organisms (K.brevis) from which they are derived (Abraham, Bourdelais, Ahmed, Serebriakov, & Baden, 2005). When K.brevis ruptures, brevetoxins are released into the water. Toxins can also be released into marine waters by excretion (Pierce et al., 2003). Whatever the mode of discharge into water, brevetoxins reach the sea surface via bubble-mediated transport from where they are aerosolized by wave action, on shore winds and breaking surf (Pierce et al., 2003). On shore winds carry aerosolized brevetoxins inland where individuals living very near or frequenting affected beaches routinely report symptoms of brevetoxin inhalation during an active red tide (Baden et al., 2005). Aerosolized brevetoxin intoxication is demonstrated by both upper respiratory symptoms such as eye and throat irritation, nasal congestion, non-

productive cough and symptoms of the lower respiratory tract like tightness of the chest, wheezing and shortness of breath. Several studies performed on persons that visited Florida beaches showed decreased respiratory function during a red tide event in contrast to periods without red tides. In all studies, participants reported upper respiratory symptoms after going to the beach and in some cases pulmonary function tests indicated that individuals who had physician diagnosed asthma had measurable respiratory impairment after exposure (Fleming et al., 2005). Because the symptoms of both ingested and aerosolized brevetoxins are only slightly hazardous as to cause discomfort in most people, it is highly likely that many of these cases are unreported. Even so, one study revealed that in a single Florida County, the annual capitalized marginal costs of hospital emergency room visits for respiratory illnesses due to K.brevis blooms ran anywhere between $0.5 to $4 million depending on bloom severity (Hoagland et al., 2009). A recent environmental health publication on the effects of aerosolized brevetoxin in mammals provides evidence showing that following exposure, brevetoxin-DNA adducts were formed in rat liver tissue. The formation of such structures likely damages DNA and RNA and may serve as an important step in the development of many types of cancer (Radwan & Ramsdell, 2008). Examination of rat tissue DNA revealed that metabolism of brevetoxins led to synthesis of highly reactive intermediates that are capable of forming nucleotide adducts (Radwan & Ramsdell, 2007). While the exact effect, if any, of the formation of these intermediates is yet to be determined in humans, Radwan and Ramsdell (2007) suggest that the evidence for adduct formation in rat mammalian cells may be a good indicator for gene toxicity in humans as well (Radwan & Ramsdell, 2007).

A joint study involving a number of organizations including the Fish and Wildlife Research Institute, Mote Marine Laboratory and NOAA, found that marine mammal mortalities may occur long after HABs are no longer visible (Flewelling et al., 2005). The study showed a significant link between the death of 34 endangered Florida manatees in Southwest Florida in 2002 and 107 bottlenose dolphins off the Florida panhandle in 2004. Extensive analysis of the contents of the mammals' stomachs and tissue tests on all affected animals indicated high concentrations of brevetoxins. Remarkably, mortalities in both cases occurred in the presence of low concentrations of K. brevis in the waters off the Florida coast. Flewelling et al. (2005) found that in manatee and dolphin mortalities, brevetoxin intoxication occurred by way of the animals' diet. High concentrations of brevetoxins were found in seagrass in the stomachs of the dead manatees while in dolphins, accumulated levels of brevetoxins were found in menhaden, the plankton-eating fish that the mammals prey on. It was determined that live fish are capable of transferring accumulated brevetoxins to higher trophic levels leaving open the possibility of human intoxication by consumption of contaminated fish (Flewelling et al., 2005). As expected, the most widely documented effects of brevetoxin poisoning are seen in the fishing industry where in the event of a HAB, marine life is decimated (Brand & Compton, 2006; Flewelling et al., 2005). Fish kills due to HABs such as those caused by K. brevis in the GoMex alone cost the seafood industry in excess of $20 million annually while total economic costs across the United States have been estimated at $82 million (NCCOS, 2009). According to a report by Central Florida News, commercial and sport fishing have more profound economic consequences on Florida

than either Disney or the state's three NFL teams combined (CFN 13, 2009). In perspective, Forbes Magazine's 2009 NFL team valuation estimated the worth of all three of these NFL teams at a combined total of $700 million (Forbes) while Florida's fishing industry contributes $6 billion to the state's economy. Justifiably Karenia brevis and particularly its unfavourable effects on both human and economic health conditions are of growing concern to industry and public health officials. This concern has translated into a significant level of research to identify the mechanisms by which K. brevis blooms are initiated and maintained. One reasonable line of research that may lead to a better understanding and ability to manage the adverse effects generated by K.brevis is to examine the molecular mechanisms underlying general cellular growth and behaviour that, in turn, determine and/or influence HAB dynamics. In particular, examining the means by which certain genes are regulated during growth and development is essential to understanding the ecology and biology of K. brevis.

Regulation of Gene Expression in *Karenia brevis*

Transcription

Gene expression is described as a two-step event that begins with transcription of DNA into messenger RNA (mRNA) followed by synthesis of protein from mRNA through a process known as translation. In the case of non-protein coding genes, however, the final product is a RNA molecule. The ultimate synthesis of proteins is controlled by efficient mechanisms that effectively turn genes on or off as needed. In most eukaryotes, the control of gene expression is primarily executed at the level of transcription. Regulation is achieved by binding of transcription factors to promoter

sequences on the DNA fragment targeted for transcription in the presence of RNA polymerase II, the enzyme that specifically catalyzes the formation of mRNA. Each gene has a promoter sequence that determines both the starting point and the direction of the transcription mechanism. A vast majority of genes contain a consensus promoter sequence, found 25-35 bases upstream of the start site, known as the TATA box (Hahn, 2004). Other cis-acting elements, i.e. sequences found on the same DNA molecule as the gene being regulated, also contribute to gene regulation in eukaryotes. Cis-acting elements include enhancers, sequences that increase the rate of transcription, and silencers, those sequences that are bound by repressors to reduce the transcription rate (Griffiths, Miller, Suzuki, Lewontin, & Gelbart, 1999). Trans-acting elements on the other hand are usually proteins that bind to regulatory sequences or other regulatory machinery. These include: 1) activators which bind to enhancers in order to activate transcription, 2) repressors which bind silencers to inhibit transcription, and 3) general transcription factors for RNA Polymerase II (TFII) which are responsible for recruitment of the RNA polymerase II enzyme. The general transcription factors, in particular, identify and bind the core promoter sequence to to initiate transcription. The first of these factors, TFIID, a protein that consists of a saddle-shaped TATA binding protein (TBP) as well as a number of TBP associated factors (TAFs), recognizes the TATA box and binds to it via this TBP subunit (Griffiths et al., 1999). It has been suggested that the TBP subunit will by itself, bind TATA box-containing promoters; however, in TATA-less promoters, TAFs are required in order for TBP subunit to "saddle" the targeted gene. Before transcription can begin, additional general transcription factors, TFIIA and TFIIB,

bind and stabilize TFIID and the TBP-TATA complex, respectively, to ensure that these factors are properly bound to the promoter. This association of TBP, TFIIA and TFIIB is referred to as the DAB complex (Hahn, 2004). Perhaps a more important function for TFFIIB is that it identifies the start site and supports the recruitment of RNA polymerase II to the core promoter. A third general transcription factor, TFIIF associates with RNA polymerase II forming a complex that then binds with the DAB-promoter complex and initiates transcription. TFIIF has two subunits, one of which has helicase activity that is capable of melting the DNA at the promoter in order to expose the coding strand. Finally, a stabilizing general factor, TFIIE recruits the final element TFIIH, a protein with nine subunits, whose kinase activity allows for clearance of the promoter by phosphorylation as well as helicase activity in order that elongation may occur (Hahn, 2004).

Aside from the TATA binding protein (TBP), the general factor TFIID contains several TBP associated factors (TAFs) that are particularly important in TATA-less promoters. Because binding of general transcription factors to the core promoter is a fundamental part of initiation, organisms such as Karenia brevis that have TATA-less promoters, must have evolved methods of initiating transcription, but they have not been extensively examined. Microarray data compiled and studied by the Van Dolah lab showed that genes involved in translation were more abundant than those of the transcriptional machinery (Lidie et al., 2005). This same study found that none of the 7001 EST sequences deposited in the GenBank database showed any homology with the transcription enzyme RNA polymerase II and only one of them exhibited similarity to the stabilizing general factor, TFIIE (Lidie et al., 2005). The

absence of a majority of identifiable components of the transcriptional machinery supports the proposal that post-transcriptional control of gene regulation is the principal mechanism for gene control in K. brevis.

Yeast, bacteria and algae respond to demanding environmental situations in a similar manner. For example, E.coli cells engage what is known as the stringent response when cells are starving for nutrients. This stringent response effectively slows down the cell's metabolic rate, turns off general transcription and translational mechanisms and turns on genes that synthesize chemicals necessary for survival (Chang, Smalley, & Conway, 2002; Raskin, Judson, & Mekalanos, 2007). An environmental stress response (ESR) analogous to that seen in E.coli is triggered in yeast and algal cells where genes coding for heat shock proteins and superoxidase dismutases are activated upon exposure to stresses (Gasch et al., 2000; Van Dolah et al., 2008). Heat shock proteins (Hsps) are named for their discovery under conditions of heightened heat exposure. Hsps are highly conserved throughout all the organisms in which they have been identified and have been found to serve a protective role by holding cellular proteins in their folded conformation, supporting the destruction of irregular proteins as well as repair of denatured ones (Miller-Morey & Van Dolah, 2004). Like Hsps, antioxidant enzymes such as superoxidase dismutases (SODs) are another class of highly conserved proteins. These enzymes prevent damage to DNA and cellular membranes that are susceptible to oxygen toxicity from the formation of reactive oxygen radicals, molecules likely to be formed in algal cells during a HAB. SODs act by breaking down the toxic superoxidases into the more tolerable forms of hydrogen peroxide and oxygen (Miller-Morey & Van Dolah, 2004). As evidenced in

other organisms, the environmental stress response in Karenia brevis is characterized by s general suspension of the transcription and translation machinery to inhibit energy consuming cellular activities like metabolism. This is coupled with up regulation of transcription of stress proteins such as Hsps and antioxidant enzymes (Miller-Morey & Van Dolah, 2004; Van Dolah et al., 2008). Using microarray analysis, Lidie and Van Dolah (2007b) showed a marked decrease in transcription of genes responsible for ribosome generation and energy acquisition. On the other hand, no evidence was found for transcription of stress related proteins e.g. heat shock proteins indicating yet again that regulation of certain cellular genes is implemented post-transcriptionally, presumably at the translational level. Unlike most eukaryotes, dinoflagellates not only lack nucleosomes, the structural units that are designed to compact the organism's DNA, but their extraordinarily dense genome is such that their nuclei exist in a unique liquid crystal state (Hackett et al., 2005). Van Dolah et al. (2008) allude to the fact that the ability to carry out nuclear functions such as transcription is complicated in K. brevis because of this intricate structure. In this event, the logical conclusion is that post-transcriptional regulation, specifically via translational mechanisms, is the best possible response to controlling gene expression in dinoflagellates.

Post-transcriptional Reglation

Regulation of translational mechanisms of gene expression is a relatively well established concept that is traditionally executed at initiation. In eukaryotes, translation involves three consecutive steps of initiation, elongation and termination. In regard to regulation, the most important of these is initiation, the rate-limiting step.

The process is characterized by binding of eukaryotic initiation factors (eIFs) to an mRNA (Biochemistry, 2006). The first of these, eIFG binds another factor, eIF4E, and together they bind the 5'cap as well as the poly-A binding protein (PABP) that is attached to the mRNA's poly-A tail. Essential associations in the translation machinery include the tertiary complex which is composed of a methionine carrying t-RNA (met-tRNA), a GTP molecule and an eIF2 protein (Biochemistry, 2006). When 80S ribosomes dissociate into the small 40S and large 60S subunits, the 40S subunit binds a series of initiation factors along with the tertiary complex to form a pre-initiation complex (PIC). This in turn associates with the mRNA at the 5'cap to give an initiation complex that scans the molecule by moving in the 5'to 3'direction. The initiation complex recognizes the start codon, AUG, by base pairing with the anticodon carried by met-tRNA. The 60S subunit rejoins the small subunit when the tertiary complex disintegrates. eIF5 and eIF5B enable the disintegration by activating hydrolysis of the GTP component, leading to liberation of eIF2-GDP and release of the remaining initiation factors (Biochemistry, 2006). At this point, elongation and termination can proceed.

It has been shown in eukaryotes that regulation at the initiation site is impeded by changes in levels of the proteins necessary for initiation to advance (Jansen, De Moor, Sussenbach, & Van Der Brande, 1995). For example, the translation initiation factor eIF4E is responsible for directing ribosomes to the 5'end of mRNA molecules in order to place the ribosomes in a position to begin translation. Levels of eIF4E therefore determine the rate at which translation can occur (Jansen et al., 1995). Other examples for possible models of gene regulation at the translational level have been

proposed. Mittag, Li, and Hastings (1998) showed that in the dinoflagellate, Gonyaulax polyedra, expression of luciferase (LCF), an enzyme involved in bioluminescence, is under translational control. During the night, LCF interacts with to allow bioluminescence, but at a certain point in the day, luciferase binding proteins (LBPs) are produced to sequester luceferin and prevent binding with LCF (Lee. Mittag, Sczekang, Morsell, & Hastings, 1992; Mittag, Eckerskorn, Strupat, & Hastings, 1997). Earlier studies using Northern blot analysis indicated that even as LBP levels fluctuated in a circadian manner due to regulated synthesis and degradation, lbp mRNA levels remained constant over time signifying control by a translational mechanism (Morse, Fritz, & Hastings, 1990; Morse, Milos, Roux, & Hastings, 1989). Like lbp mRNA molecules that encode LBP, lcf mRNA levels remain constant during the circadian cycle suggesting that similarly, LCF expression is under translational control. (Mittag et al, 1998; Morse et al., 1989).

Generally, mRNA strands are extremely unstable molecules whose half-life can be improved by the presence of various factors such as a poly-A tail at the 3'end of each fragment (Jansen et al., 1995). Consequently, the half-life of an mRNA molecule can be determined by how fast it loses its poly-A tail (Jansen et al., 1995). On the other hand, some eukaryotic mRNA molecules such as Xenopus serum albumin have the innate ability to control the length of their poly-A tails (Gupta, Gu, Chernokalsakya, Gao, & Schoenberg, 1998). Xenopus serum albumin maintains the length of its short poly-A tail during degradation and this manipulation provides for an additional mode of translational control. Recently, novel mechanisms involving small RNAs (sRNA) have been described in bacteria that regulate translation. Non-

coding RNAs are capable of regulating translation of target mRNAs to repress protein levels by base pairing with the mRNAs. Conversely, a small RNA, RhyB, was found to stimulate translation of the shiA mRNA in Escherichia coli (Prevost, Salvail, Desnoyers, Jacques, & Masse, 2007). Translation of shikimate permease, an enzyme encoded by the shiA gene in E.coli, is complicated by a structure that blocks initiation of translation at the ribosome binding site (RBS). RhyB overcomes this obstruction and promotes translation by base pairing with the 5'UTR of the shiA mRNA to prevent formation of the inhibitory structure (Prevost et al., 2007). Similar sRNAs have also been found in eukaryotes (Jacquier, 2009; Vogel et al., 2003). Gottesman (2005) hypothesizes that the function of base pairing of sRNAs with a target mRNA is determined by the precision of the pairing mechanism: imprecise pairing of sRNAs at the 3'UTR of the target mRNA is a precursor to inhibition of the translational machinery whereas exact pairing permits mRNA degradation.

Investigations of the K. brevis transcriptome (Pirooznia et al., in press) revealed the presence of non-coding RNA molecules (ncRNA) whose sequences were complementary to certain mRNAs. Because of this complementarity, it is hypothesized that the ncRNA molecules and the respective mRNAs form a double stranded RNA (dsRNA) complexes, which serves to inhibit/obstruct translation or stabilizes the mRNA and enhances translation, possibly depending on the physiologic condition of the cell. The presence of ncRNA transcripts not only suggests possible models for translational controls of gene expression but also indicates that translational control may play a more important role in regulating gene expression in dinoflagellates than has been previously observed in other eukaryotes. The objective

of this thesis project is to verify the presence of ncRNA transcripts in the transcriptome and produce evidence of a possible function(s).

Anti-sense RNA

Non-coding RNAs ranging in size from 20 bp to > 200 bp have long since been known to be naturally occurring in prokaryotes and eukaryotes. Instead of encoding for proteins, ncRNAs produce RNA molecules such as rRNAs and tRNAs that play essential functions in the cell. Small regulatory ncRNAs such as microRNAs (miRNAs), small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) are also well documented examples of this extensive category. These ncRNAs play important roles in post-transcriptional regulation, mRNA splicing, rRNA modification, and RNA editing, respectively. Below, I present some of the best-studied examples of anti-sense RNAs (asRNAs) in the model organisms Mus musculus and Caenorhabditis elegans including the functional role the asRNAs play in each organism. We draw our hypothesis for the role of asRNAs in K. brevis from the analogous roles of these molecules in the above model organisms, given the feasible mechanism for sequence complementarity.

Non-coding RNA transcripts have been found for various imprinted genes in M. musculus, suggesting that ncRNA plays an important role in genomic imprinting. The mouse insulin-like growth factor 2 receptor gene, Igf2r, is expressed only from the maternal chromosome since the maternal Igf2r makes only the sense transcript. In the maternal copy of Igf2r, there is an unmethylated upstream promoter which enables transcription of the Igf2r mRNA. The second intron of Igf2r contains an imprinting element, a 2kb CpG island known as region 2. This region, located on the

downstream promoter of the maternal Igf2r is methylated. Unmethylated region 2 is critical to the formation of the antisense transcript but because this region is methylated in the maternal chromosome, asRNA cannot be synthesized. On the other hand, the paternal copy of the Igf2r gene is methylated at the promoter region, inhibiting its ability to transcribe the sense strand. However, the downstream promoter is unmethylated enabling transcription of the antisense strand. It has been suggested that the role of the ncRNA transcript in genomic imprinting is to recruit components of the transcriptional silencing machinery.

In the nematode, Caenorhabditis elegans, the timing and progression of postembryonic development is controlled by the levels of Lin-14 protein. lin-4 transcripts are regulatory, 22-nucleotide ncRNA that are complementary to the 3' untranslated region (UTR) of lin-14 mRNAs, and when bound, lin-4 inhibits accumulation of Lin-14 protein. Using C. elegans mutants lacking the lin-4 gene, researchers discovered that over time, levels of Lin-14 did not decrease which correlated with the effect that these mutants repeat the larval stage and fail to proceed to later developmental stages (Eddy, 2001). Since lin-14 transcript levels were found to be constant throughout the development of the C. elegans mutants, it was determined that lin-14 is negatively regulated post-transcriptionally. By hybridizing with the 3'-UTR of lin-14 RNA, lin-4 transcripts interfere with translation of *lin-14* effectively down regulating its expression (Eddy, 2001).

Figure 2 depicts four of several genes from the McLean EST library (Pirooznia et al., in press) that indicated the presence of one or more asRNAs. For all four genes there exist regions of variable length that possess perfect complementarity

between the respective mRNA and asRNA sequences. These double stranded regions vary in size (860 bp for actin, 745bp for ADP-ribosylation factor, and 240 bp for the mitochondrial gene cytochrome oxidase I). Cytochrome B has regions of complementarity with two different asRNAs, one measuring 470 bp in length and the other 420 bp. This EST data indicated that the asRNAs have unique 3' ends as evidenced by their poly-A tails and the potential for unique 5' ends that will be confirmed through RACE-PCR experiments.

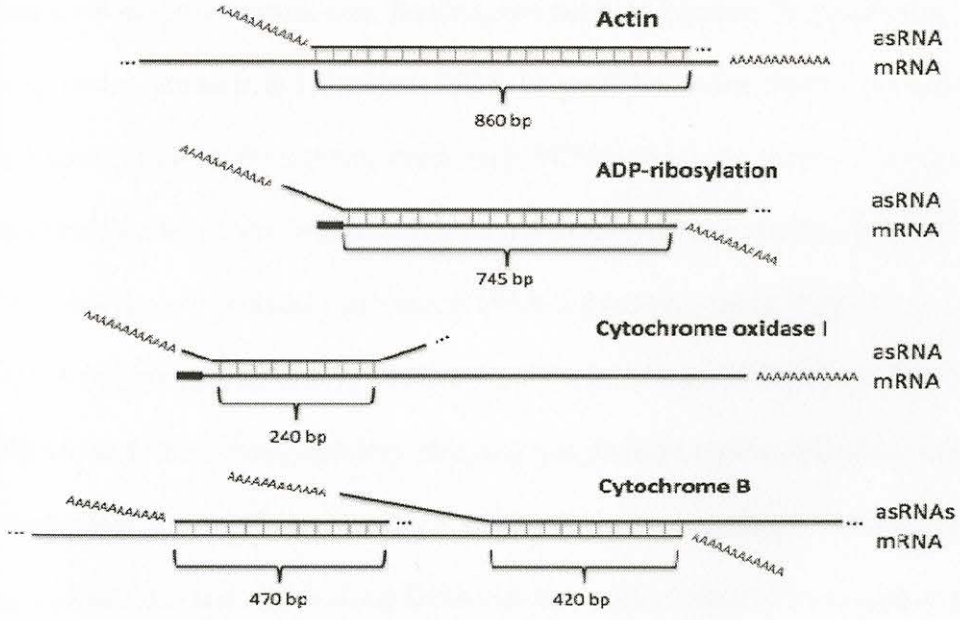


Figure 2. Schematic Representations of Complementary mRNA and asRNAs. Representative mRNA and asRNA sequences are shown from contigs in the McLean K. brevis EST database that were annotated as labeled. Regions of perfect complementarity are bracketed. Triple dots indicate that the sequence may or may not continue in the specified direction—in many cases, the ESTs do not cover the full-length of the transcripts, and 5' ends in particular are missing. The bold line represents the spliced leader sequence denoting the true 5' end of the mRNAs. In addition to the genes listed, the sense and antisense transcripts for Photosystem D2 form a construct similar to that depicted by Actin, with a 530 bp overlap, and the sense and antisense cytochrome oxidase 3 transcripts binds in a similar manner to that shown for cytochrome oxidase 1 with a 360 bp overlap.

CHAPTER III

EXPERIMENTAL DESIGN AND METHODOLOGY

Genes Under Study

Analysis of the McLean EST library indicated the presence of one or more asRNAs for certain genes. The following genes are examples of some of these genes

Nuclear Genes:

PCNA. Proliferating Cell Nuclear Antigen (PCNA) is a nuclear protein that has been found to play several roles in important cellular events including DNA replication, DNA repair, coordinating synthesis of Okazaki fragments and cell cycle regulation (Jonsson & Hubscher, 2005; Maga & Hubscher, 2003). An essential component of the replication machinery, PCNA serves an ancillary function to DNA polymerase by increasing the processivity of this enzyme. In fact, PCNA is often referred to as an auxiliary protein of DNA polymerase-delta. Together with another DNA polymerase accessory protein, Replication Factor C (RFC), PCNA and DNA pol-delta form a clamp-loading complex that ensures rapid and precise DNA replication. PCNA forms the DNA sliding clamp, a ring-shaped structure that not only encircles and slides along DNA but also tethers DNA polymerase to the DNA helix preventing it from falling off the template hence encouraging synthesis of long strands (Miyata et al., 2004; Tsurimoto, 1999). Because of its significance in DNA replication, it follows that PCNA is a vital component of the cell cycle and could even be useful as a growth marker (Baserga, 1991; Brunelle & Van Dolah, 2011; Zhao et al., 2009).

Quantitative PCR, Western blotting and immunolocalization studies performed on a number of S-phase genes by Brunelle and Van Dolah (2011) indicated that while PCNA transcripts were constitutively expressed throughout the cell cycle, including the S-phase, their corresponding proteins showed differential expression with maximal protein abundance in the S-phase. Given the level expression of PCNA transcripts, these results not only provide support for a post-transcriptional model of gene expression in Karenia brevis, they also reveal the use of PCNA as a viable reference gene in K. brevis qPCR studies. Western blotting results gave additional support for gene control at the translational level as PCNA bands showed an increase in molecular weight from 28-kDa to 37-kDa during S-phase. Ubiquitination of PCNA by addition of the 8.5-kDa ubiquitin molecule or sumoylation by addition of a 9-kDa small ubiquitin-like modifier (SUMO) are examples of post-translational modifications identified in other eukaryotes and may account for the increase in protein size in K. brevis (Brunelle & Van Dolah, 2011). Based on results obtained from Brunelle and Van Dolah's qPCR experiments signifying that expression levels of PCNA remained unchanged under light and dark conditions, we used PCNA as a reference gene in our qPCR studies.

Actin. Actin is a highly conserved protein found in virtually all eukaryotic cells where it is crucial in filament formation. Actin filaments are a major component of the cytoskeleton which is involved in several cellular processes including motility of cells and maintaining cell shape (Harata, Karwan, & Wintersberger, 1994; Reisler & Egelman, 2007; Zheng, Han, Bernier, & Wen, 2009). In dinoflagellates, actin has previously been detected in both the cytoplasm and the nucleus by

immunocytochemistry and observed using light and electron microscopy (Perret, Albert, Besseau, & Soyer-Gobillard, 1993; Soyer-Gobillard, Ausseil, & Geraud, 1996). More specifically, a number of papers have been published indicating the presence of actin in *Karenia brevis*. For example, analysis of an expressed sequence tag (EST) library created by Lidie et al. (2005), revealed 15 copies of the cytoskeletal protein. Four unique actin genes (act1-4) have been found in *K. brevis* (Lidie et al., 2005; Saldarriaga, McEwan, Fast, Taylor, & Keeling, 2003). The GenBank sequence that showed maximum identity with the actin found in the McLean EST libraries was *Karenia brevis* actin (act2).

Evidence exists supporting the function of actin in the formation of the cleavage furrow of dividing cells in the dinoflagellate Crypthecodinium cohnii and Prorocentrum micans as well as several ciliates and yeasts (Harata et al, 1994; Perret et al., 1993; Soyer-Gobillard et al., 1996), providing yet another mechanical function for actin. Furthermore, in several other eukaryotes, nuclear actin has been shown to be of critical importance in the transcription machinery (Dominguez & Holmes, 2011; Zheng et al., 2009), and suggestions have been made indicating that nuclear actin may play a role in compacting the nucleolus in certain mammalian cells (Funaki, Katsumoto, & Iino, 1995).

Previous research has shown evidence of the abundance of actin in the dinoflagellate transcriptome (Bachvaroff & Place, 2008) and multiple actin copies have been specifically identified in Karenia brevis (Lidie et al., 2005; Saldarriga et al., 2003). Additionally, Bachvaroff and Place (2008) were able to correlate genes found in tandem repeats, such as actin, with high levels of expression. Almost all

highly expressed genes, including actin, were trans-spliced. These results suggest that trans-splicing may be critical in controlling expression as for the most part only those genes with the spliced leader are translated (Bachvaroff & Place, 2008; Lidie et al., 2005).

ADP-ribosylation factor (ADP-Ribo). ADP-ribosylation factors (Arf) are part of a family of GTP-binding proteins involved in intracellular trafficking, a process that entails delivery of proteins to their respective targets after translation (Lee et al., 2002). The intracellular trafficking mechanism has been well studied in animal, plant and yeast cells (Lee et al., 2002), but like many other cellular processes, little data exists elucidating the mechanism of signal transduction in dinoflagellates. In general, transport between organelles, e.g., from the endoplasmic reticulum to the golgi, is mediated by vesicles. Formation of these vesicles is an intricate process that is regulated by a number of protein families including members of the Arf family (Donaldson & Segev, 2009). Gene expression studies on the dinoflagellate Alexandrium catenella revealed ADP-ribosylation factors involved in vesicle transport and sequencing data for the dinoflagellate Pyrocystis lunula from Okamoto and Hastings (2003) showed a molecule with similarity to a trafficking protein, ADP-ribosylation factor.

Putative Chloroplast Genes

Photosystem II D2. The photosystem II reaction center protein D2 (PsbD or photo D2) is encoded by psbD mRNA (Nickelsen, Fleischmann, Boudreau, Rahire, & Rochaix, 1999). As a result of endosymbiotic gene transfer, a majority of genes previously encoded in the plastid genome were transferred to the nucleus of peridinin-

containing dinoflagellates. The exception includes genes that encode core proteins of the photosystem such as PsbD (Hackett et al., 2004; Yoon et al., 2005). Results based on comparative analysis of *Karenia brevis* putative plastid genes including psbD indicated that these genes have structural similarities (nucleotide content) with genes known to be localized in the chloroplast genome in peridinin-containing dinoflagellates (Yoon et al., 2005). Photosystem II (PSII) is contained within the thylakoid membranes of the chloroplast where it absorbs light at a wavelength of 680nm. It then uses this light energy to oxidize water molecules to obtain electrons and liberate oxygen (Melis, 1999; Miles, 2003). Found within the core complex of PSII are the pigments, chlorophyll a and β-carotene, which are bound to the light harvesting or antenna proteins, CP43 and CP47. These pigments absorb light energy and pass it on to the reaction center proteins D1 and D2 which bind P680, a pair of chlorophylls that absorb light at a wavelength of 680nm. Energy absorbed from these photons excites P680 converting it into a reducing agent that transfers electrons from one electron acceptor to the next before they reach the next complex in the electron transport system (Barber, 2002; McClean & Johnson, n.d.; Melis, 1999; Miles, 2003).

Previous research has shown that D2 is involved in synthesis of the light harvesting protein CP47 and the PSII subunit D1 (Komenda et al., 2004). Given the essential roles that both CP47 and D1 play in PSII, Komenda et al., (2004) concluded that D2 protein is of critical importance in the assembly of the PSII reaction center complex in Synechocystis. The function of PSII proteins in the process of photosynthesis is extensively studied in plant, algal and cyanobacterial species (Barber, 2002; Komenda et al., 2004), yet beyond its use in phylogenetic and

evolutionary studies (Yoon et al., 2005), very little data exists regarding other roles this protein may play in Karenia brevis. For our purposes, we chose to use Photosystem II D2 because as a chloroplast gene, it is undoubtedly involved in photosynthesis. As a result, it is a fitting candidate for gene studies investigating differential expression under day and night conditions.

RuBisCO. Ribulose-1,5-biphosphate carboxylase oxygenase (RuBisCO) is thought to be the most abundant protein on earth (Cooper, 2000; Dhingra et al., 2004; Palmer, 1996). In fact, Casper, Paul, Smith, and Gray (2004) were able to detect Karenia brevis cells in the environment using Rubisco mRNA (rbcL). The choice of Rubisco as a target was influenced by the high expression levels for rbcL mRNA (Casper et al., 2004). Rubisco is of critical importance because it catalyzes the carbon fixation reaction between carbon dioxide and ribulose-1,5-biphosphate, the first step in the Calvin cycle (Cooper, 2000). Rubisco is found in two distinct forms. Form I, the typical type of Rubisco, is a protein made up of both small and large subunits, is encoded in the chloroplast, and is commonly found in both prokaryotic and eukaryotic organisms. Form II was identified for the first time in peridinin-dinoflagellate chloroplasts where it was found to be nuclear-encoded and composed of only large subunits (Morse, Salois, Markovic, & Hastings, 1995; Palmer, 1996; Yoon et al., 2005). Analysis of K. brevis ESTs to determine the fate of nuclear-encoded plastid genes that had previously been detected in a peridinin-dinoflagellate, revealed the absence of form II Rubisco (Yoon et al., 2005). This suggests a return to plastid-encoded form I Rubisco in fucoxanthin dinoflagellates, especially considering that form I Rubisco has been successfully isolated in Karenia brevis (Yoon et al.,

2005). Predictably, we selected Rubisco for the abundance of its mRNA, a feature that facilitates molecular investigations.

RNA Processing

RNA for strand-specific PCR, RT-qPCR and RACE reactions was harvested using either filtration with a vacuum pump on a 0.8μm- 3.0μm Versapor® Acrylic Copolymer Membrane Disc Filter (Pall Corporation, USA) or centrifugation in a centrifuge at 1500rpm for five minutes when time was limited. Even though cells were spun down for only five minutes during centrifugation, vacuum filtration was the method of choice for the majority of RNA extractions. Unlike filtration, centrifugation is likely to induce a stress response in the cells which would alter the transcriptome (Neymotin, 2011). To obtain total RNA, cell pellets were treated using TRIzol® reagent (Invitrogen, Carlsbad, USA). TRIzol reagent yields higher concentrations of RNA than column-based extraction such as Qiagen's RNeasy kit (Santiago-Vazquez, Ranzer, & Kerr, 2006); however, since the TRIzol extraction procedure is relatively lengthy, I did use the RNeasy kit when time was an issue.

Two hundred milliliters of relatively dense Karenia brevis culture were collected on the filter paper and washed off with 1 ml TRIzol reagent into a polypropylene tube. The mixture was homogenized by pipetting to allow for proper cell lysis before proceeding to phase separation with acidic phenol/chloroform. The upper aqueous phase containing RNA was pulled off and precipitated with 100% isopropanol. The pellet was washed with 75% ethanol and resuspended in DEPC-treated water. RNA concentration and integrity was determined by spectrophotometric analysis using a NanoDrop ND-2000 (Thermo Fisher Scientific).

Additional assessment for purity of RNA used in RT-qPCR samples was done using Agilent's BioAnalyzer 2100. To eliminate any residual DNA, all RNA samples were DNase treated with RQ1 DNase (Promega, Madison, WI) for forty minutes at 37°C before being added to a Reverse Transcription reaction. When RNA was not used immediately after extraction, it was stored at -20°C in 1/10 volume 3M sodium acetate, pH 7.5 and 2.5 volumes of 100% ethanol.

Reverse transcription was performed for two hours at 42°C and then heated to 95°C for ten minutes. The reaction mixture contained: reverse transcriptase and SUPERaseIn, an RNase inhibitor (Life Technologies, New York, USA), 2.5mM dNTPs as well as a gene/strand specific primer. A volume of the newly synthesized cDNA was then added directly to a PCR mastermix containing 25mM dNTPs, 0.5uM reverse and forward sequence-specific primers, 10X PCR buffer, Taq polymerase and DEPC-treated water to a final reaction volume of 20µl. Sample tubes were loaded into a thermocycler and run at variable annealing temperatures (dependent on primer melting temperature). The following PCR parameters were set for our Strand-Specific and RT-PCR reactions (see Table 1).

Table 1

Strand-Specific and RT-PCR Parameters

	Temperature	Time (minutes)	Repeat(# times)
Hot Start:	94°C	3	
Denaturation:	94°C	0.5	30X
Annealing:	X °C	0.5	30X
Elongation:	72°C	2	30X
Final Elongation:	72°C	10	

Strand-specific Polymerase Chain Reaction

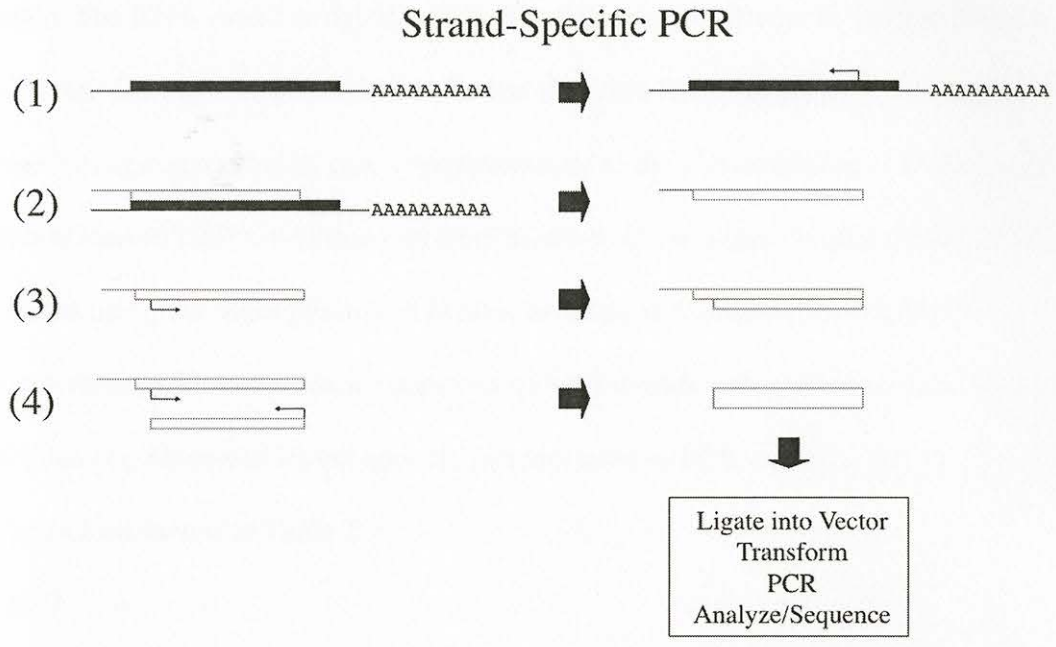


Figure 3. Steps of a Strand-specific PCR. Step (1): reverse transcription of mRNA into ssDNA and formation of mRNA-DNA hybrid with reverse transcriptase. A single gene/strand specific primer was used for each desired gene and asRNA or mRNA strand. Step (2): degradation of mRNA with RNAse H retaining only cDNA representing target molecule. Steps (3+4) formation of dsDNA during PCR with the same primer used during reverse transcription along with an additional gene/strand specifc primer. The amplified products can be run on a gel for analysis and used in further downstream applications.

A set of experiments to verify the presence of the non-coding RNAs was performed. Strand-specific PCR is a molecular technique that involves the generation of a double stranded DNA (dsDNA) molecule from the reverse transcription of a single-stranded RNA molecule. All RNA was treated with DNase prior to reverse transcription. Figure 3 demonstrates how the procedure converts an RNA molecule to single-stranded complementary DNA (cDNA) by reverse transcription. The reverse transcription reaction is primed by a single gene-specific primer (GSP) that recognizes either the sense or anti-sense RNA for a particular gene of interest. Other

added reagents include 25mM dNTPs, a reverse transcription buffer and exposure to the RNase Inhibitor, Superase-In, to minimize any chance of degradation of the RNA sample. The RNA strand in the hybrid is then degraded by RNase H, leaving behind the cDNA. During a PCR reaction, a double stranded DNA fragment (dsDNA) is formed using a second GSP now complementary to the cDNA product. PCR using the first and second GSP's will result in amplification of the target molecule. Concurrent reactions using the same primers in reverse order or new primers specific for the reverse strand will determine if either one or both strands are present in the extracted RNA sample. Gene and strand specific primers used in PCR reactions for all 8 genes are presented below in Table 2.

Table 2

Primer Sets and Amplicon Sizes

PRIMER SETS	PRODUCT SIZE (bp)
Actin RO (818) and FO (622)	196
ADP FSL and FSR	241
PCNA R and F	164
PhotoD2 RO and FO	335
Rubisco RO and FO	263
CytoB set 1 (921L and 921 R)	224
CytoB set 2 (FL and FR)	169
CoxIII CytoC3 RS and CytoC3 RR	178

Negative controls were set up to discount any potential contamination concerns for asRNA and mRNA samples for all genes tested. In all the RNase pre-treatment control samples, RNA was treated with a mixture of RNases to degrade all RNA molecules. In the negative reverse transcriptase controls (RT-), reverse transcriptase, the key ingredient for cDNA synthesis was excluded. A negative control using reverse transcriptase (RT+) in RNase pre-treated samples was also included. All these controls check for DNA contamination, and the absence of specific bands in all the negative controls is indicative of a DNA-free sample. mRNA samples were reverse transcribed with the appropriate reverse primer while asRNA samples were reverse transcribed using a forward primer. During PCR, cDNA generated from both transcripts was amplified using reverse and forward primers. After PCR, samples of each reaction were mixed with GelStar (Lonza Bioscience, Switzerland) and run on 1.5% agarose gels. Gel images showing the presence or absence of reaction-specific PCR bands were captured on a BioRad Gel-Doc system (BioRad, Hercules, CA).

Reverse Transcription-PCR (RT-PCR)

RT-PCR works much like strand-specific PCR in that a gene specific primer is used to generate cDNA from a specific RNA transcript. The reverse transcription step is performed using DNA-free RNA, reverse transcriptase enzyme, a gene-specific reverse primer, RNase-inhibitor and reverse transcriptase buffer. RT was done with a gene specific primer (GSP) or a poly-T primer for 5' and 3' ends, respectively. Unlike strand-specific PCR, RNase H is not used to degrade the RNA template in RT-PCR. After cDNA synthesis, a spliced leader primer or a different GSP was used to amplify

5' or 3' ends during PCR. Because of some uncertainty regarding whether or not the ends of the asRNA transcripts contain the spliced leader sequence, we used RACE-PCR to capture both ends of the asRNA sequences.

Rapid Amplification of cDNA Ends (RACE)

Strand-specific PCR is a molecular tool that uses two oligonucleotide primers, one that is specific to the coding strand and the other to the non-coding strand, to amplify DNA templates (Frohman, Dush, & Martin, 1988). Still, this technique is limited to characterization and amplification of regions of known sequences. Additionally, because reverse transcriptase rarely generates complete sequences, there's a poor representation of 5' ends of genes in cDNA populations (Clontech). To provide further evidence of asRNAs in the Karenia brevis transcriptome, we attempted to capture full length mRNA and asRNA sequences for actin, photo D2 and ADP-Ribo genes using rapid amplification of cDNA ends (RACE). In RACE-PCR, first strand cDNA synthesis using a primer for a small known sequence within a transcript (derived from the EST data) is followed by PCR (using the same GSP and the Universal primer from the RACE kit) to obtain the complete sequence of the ends of the RNA transcripts.

We used the SMARTer RACE cDNA Amplification kit provided by Clontech for both 5' RACE (mRNA and asRNA) and 3' RACE (mRNA and asRNA) reactions. cDNAs are synthesized during reverse transcription using both SMARTer II A oligonucleotide and a type of Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase, SMARTScribe RT, which displays both terminal transferase and template switching activities. When the RT reaches the end of the mRNA molecule, it

adds several non-template residues to the 3' end of the cDNA strand, an effect of its terminal transferase activity. The SMARTer oligo II A then anneals to the newly added stretch of cDNA and serves as a template for the RT enzyme, which switches templates to the SMARTer oligo from the mRNA. SMARTScribe RT switches templates only after it reaches the end of the mRNA strand and as a result, the SMARTer oligo can only be incorporated into full-length cDNAs. In 3' RACE, cDNA synthesis of the first strand is primed with an oligo(dT) primer, 3'-RACE CDS Primer A, which positions itself at the start of the poly A-tail. This CDS primer contains two degenerate nucleotides at the 3' end of the mRNA that ensure initiation of cDNA synthesis only at this site. After first strand synthesis during reverse transcription, both 5' and 3' RACE-Ready cDNAs contain the SMARTer sequence. Therefore, the same primer, Universal Primer A Mix (UPM) can be used for both RACE-PCR reactions. UPM is complementary to the SMARTer sequence and can be used along with gene specific primers that recognize a small known sequence within both 5' and 3' RACE cDNAs.

Using the SMARTer RACE Amplification kit, unknown regions of the 5' and 3' ends of actin, ADP-Ribo, and photo D2 were amplified. One microgram of Dnase-treated total RNA extracted via TRIZOL reagent was prepared for reverse transcription. For the 5' RACE reaction, RNA and 5' CDS Primer A mixture was incubated at 72°C for three minutes then cooled to 42°C for two minutes before the SMARTer IIA oligonucleotide was added. The same procedure was repeated for the 3' RACE reaction but without addition of the SMARTer IIA oligo. A mastermix containing SMARTscribe RT, RNase Inhibitor and a buffer mix (5X First-strand

buffer, DTT and dNTP Mix) was then added to both 5' and 3' RACE reactions. For reverse transcription to proceed, reactions were incubated at 42°C for ninety minutes and then heated to 70°C for ten minutes in a thermocycler. The first-stand reaction products were diluted in Tricine-EDTA buffer before proceeding to PCR amplification. For first round amplification, genes were amplified with the universal primer (UPM) and a gene specific primer (GSP). A second round of amplification was performed using the diluted product from round one amplification, a nested universal primer (NUP) and nested gene specific primers or inner primers (NGSP). Primers used in both amplification rounds for all genes are shown in Table 3.

Table 3

Primer Sets Used in Each Round of RACE PCR

Gene	First Round PCR	Second Round PCR
Actin 5' mRNA	UPM and Actin RO	NUP and Actin RI
Actin 5' asRNA	UPM and Actin FO	NUP and Actin FI
Actin 3' mRNA	UPM and Actin FO	NUP and Actin FI
Actin 3' asRNA	UPM and Actin RO	NUP and Actin RI
Photo D2 5' mRNA	UPM and D2-RO	NUP and D2-RI
Photo D2 5' asRNA	UPM and D2-FO	NUP and D2-FI
Photo D2 3' mRNA	UPM and D2-FO	NUP and D2-FI
Photo D2 3' asRNA	UPM and D2-RO	NUP and D2-RI
ADP-ribo 5' mRNA	UPM and ADP-RO	NUP and ADP-RI
ADP-ribo 5' asRNA	UPM and ADP-FO	NUP and ADP-FI
ADP-ribo 3' mRNA	UPM and ADP-FO	NUP and ADP-FI
ADP-ribo 3'mRNA	UPM and ADP-RO	NUP and ADP-RI
ADP-ribo 3' asRNA	UPM and ADP-RO	NUP and ADP-RI

RT/RACE PCR products were run on a 1.5% low melt agarose gel. Bands representing each amplified product were excised from the gel, and DNA was extracted using Qiagen's QIAquick Gel Extraction kit (Qiagen, Valencia, CA). DNA integrity and concentration was determined by NanoDrop before incubation with pGEM-T Vector (Promega, Madison, WI) in a one hour room temperature ligation reaction. We followed the manufacturer provided ligation protocol using 5ul 2X Rapid Ligation Buffer, 1ul 50ng pGEM-T Vector, purified DNA, 1 µl T4 DNA Ligase and added water to a final volume of 10 µl. The concentration of DNA to be added was determined using the calculation below:

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times (\text{insert:vector molar ratio}) = \text{ng of insert}$$

The ligated samples were then added to Z-competent E.coli cells (Zymo Research, Irvine, CA) and the mixture was incubated overnight for transformation. The Z-competent cells have a selectable marker for carbenicillin resistance, ensuring that only those cells with the vector will grow. Transformed cells were then picked into two ml LB agar with carbenicillin and incubated overnight with shaking at ~220rpm. To extract and purify plasmid DNA, the cultured cells were processed via plasmid miniprep using Zyppy plasmid miniprep kit (Zymo Research, Irvine, CA). Purified plasmid DNA was then sent off to be sequenced by Eurofins MWG Operon (Huntsville, AL). Data was compiled and analyzed to identify and characterize full length sequences for ADP-ribo, photo D2 and actin genes.

Immunofluorescence

Since we expect that complementary asRNA and mRNA molecules to form double stranded RNA, it is important to show that these double stranded RNA

molecules do in fact exist in vivo. We used an indirect immunofluorescence (IF) technique that involves using labelled antibodies specific for double stranded RNA molecules in order to detect the presence of the dsRNAs.

One milliliter of stock glutaraldehyde (25%) was added to a nine milliliter culture of K. brevis in Corning tube. The cells were gently mixed, and left to sit at room temperature for twenty minutes to allow for fixation. Glutaraldehyde is a cross-linking reagent which preserves cell structure and allows the antibody to have easy access to its specific antigen. The glutaraldehyde was prepared with conditioned media to create a solution that is isotonic to the cells. The fixed sample was divided into four separate 1.5ml microfuge tubes which were all spun down at 1000rcf for five minutes to pellet the fragile K. brevis cells without rupturing them. Pellets were washed in room temperature phosphate buffered saline (PBS) prior to blocking for thirty minutes at 4°C with 3 % bovine serum albumin (BSA) in PBS. After blocking, samples were washed in chilled PBS and then treated as follows: 1) a sample was incubated with the primary antibody only; 2) a second sample was incubated both the primary antibody and the secondary antibody, an AMCA-labelled goat anti-mouse (Vector Labs), 3) a sample pre-treated with RNase A, and 4) a sample pre-treated with RNase III. AMCA is a fluorescent dye bound to the secondary antibody and will emit signal in the blue range (~ 450nm) upon excitation (350nm) under a confocal fluorescence microscope. The two pre-treated samples along with the first sample (test sample) were incubated with both the secondary and primary antibodies. The primary antibody used was the mouse J2 monoclonal antibody (Scions, Hungary). After the relevant incubations with RNases and/or antibodies, the cells were washed

and then counterstained with propodium iodide to visualize the nuclei. A fifth sample was introduced in the initial experiment to optimize staining with propidium iodide. In this experiment, cells were stained with only propidium iodide. A Zeiss 510 Meta confocal laser-scanning microscope was used to detect and visualize the dsRNAs.

Quantitative Real-time PCR (qPCR)

Traditional or conventional PCR methods have been and continue to be beneficial in determining the presence or absence of a target molecule and the quality of DNA amplicons including primer dimer contamination. However, this qualitative analysis can only be performed on an agarose gel after the PCR reaction is completed, hence the name, end-point PCR (Smith & Osborn, 2008). Quantitative PCR (qPCR) provides a method that determines the amount of DNA generated in each cycle. In a qPCR run, an intercalating fluorescent dye that binds to double stranded molecules is added to a PCR sample and as the reaction proceeds, changes in fluorescence levels are detected and recorded. Changes in fluorescence are proportional to the changes in DNA concentration as amplification proceeds. Predictably, qPCR is now the most preferred method for quantification of nucleic acids while conventional PCR continues to be used for qualitative analysis (Bar, Kubista, & Tichopad, 2011; Pfaffl, 2001; Smith & Osborn, 2008; Szajek, 2007). Depending on the purpose of the experiment, absolute or relative quantification procedures are followed. Absolute quantification is used to determine the number of copies of a target gene that are present in a particular sample. On the other hand, relative quantification is useful for comparing the relative amounts of nucleic acid present in equal amounts of different samples (BioRad, 2012).

For our purposes, we used the relative quantification method to compare gene expression levels in night versus day samples for actin and photo D2 genes using PCNA as the reference gene. Reference genes are genes whose mRNA is relatively stable under all conditions and can therefore be reliably used as a normalizer in comparing expression levels of other genes. For each day and night reaction, RNA was extracted at 9AM and 9PM, respectively. DNase treatment and reverse transcription was performed in the exact same manner for both RNA samples to prevent biases or inconsistencies that may be present in the results. A standard curve was prepared for each gene using a DNase-treated RNA sample. cDNA for the curves was generated from day mRNA samples (untreated) and amplified by PCR. The PCR product was purified using the NucleoTrap PCR purification Kit (Machery-Nagel, Bethlehem, PA), eluted in 25µl elution buffer and DNA concentration determined by NanoDrop. The number of copies/µl of DNA was calculated using an online calculator developed by Andrew Staroscik (URI Genomics and Sequencing Center). Together with the size of the gene, this DNA copy number was used to determine the volume required to create the initial point in the standard curve. Five 10-fold dilutions were then created from this starting concentration and added to a PCR master mix containing nuclease-free water, 0.5µM of the appropriate forward and reverse primers and iQ SYBR-Green Supermix (BioRad, Hercules, CA). The SYBR-Green Supermix contains DNA polymerase, buffer solution, fluorescein dye and dNTPs. For sample preparation, cDNA created from a starting material of 1.5µg of DNA-free RNA was added to a separate mastermix identical to that described above. A separate mastermix was prepared for each test and control sample, and all samples were

performed in triplicate. Standard curve and target gene samples were loaded onto a qPCR machine (BioRad CFX 9000) along with no-template (NTC) and no-RT (NRT) controls and run at an annealing temperature of 58°C.

In summary, each qPCR run consisted of two standard curves, one for the reference gene, PCNA, and the other for the target gene, actin or photo D2. Also included for each reference and target gene were the test sample (asRNA) and the calibrator sample (mRNA) for both day and night reactions. To validate the efficiency of our assay and to generate reliable data, negative controls (NTC and NRT) were also added to the run. At the end of each run, a software-generated gene expression chart is produced showing normalized fold expression levels as well as a melt curve confirming that only the targeted product was amplified. Calculations to determine the expression level of one gene relative to another were performed manually using the ΔΔ Ct method. For example, for both the test (asRNA) and calibrator (mRNA) samples, the Ct value of the target gene is normalized to that of the reference gene to get the Δ Ct value. The Ct of the test sample is then normalized to that of the calibrator to attain the ΔΔ Ct value. The expression ratio is then determined using the formula 2 e (-ΔΔ Ct).

CHAPTER IV

ANALYSIS OF DATA

Strand-specific Polymerase Chain Reaction

The McLean EST dataset contains numerous examples of EST clusters that contain at least one sense and one antisense transcript (Pirooznia et al., in press). The sequence information in selected representatives from this subset of genes was used to create gene-specific/strand-specific primers. The genes were mostly annotated and ranged over all three genomes within the cells (i.e. nuclear, chloroplast, and mitochondrial). Gel electrophoresis analyses of strand specific-PCR products verified the presence of both asRNA and mRNA transcripts for tested nuclear, mitochondrial, and putative chloroplast genes (Figure 4). (For the following discussion of the results, lane designations apply to all gel images except "CytoB set 2" and "CoxIII" where the ladder was added before the test samples. For those two gels, the lane designations are in the same order but are +1 relative to all other gels.) Lanes 1, 2, 5, and 6 were PCRs performed on samples that were reverse transcribed using reverse primers (R) that are specific for mRNA transcripts. Lane 3, 4, 7, and 8 were PCRs performed on samples that were reverse transcribed using forward primers (F) that are specific for asRNA transcripts. For lanes 1-4, the negative RNase PT controls, total RNA was pre-treated with a mixture of RNases prior to the reverse transcription reaction. In all odd lanes, the RT negative (RT-) controls, the enzyme reverse transcriptase was not added to RNA samples before proceeding with reverse transcription. As expected, no specific product was shown in the negative controls for any of the genes. In total, the absence of specific PCR bands in lanes 1-4, 5, and 7 for

each gene tested supports the conclusion that the presence of an asRNA- or mRNA-specific PCR product was due to reverse transcribing the relevant transcript and not from contaminating DNA or RNA that is hanging around after the reverse transcription reaction. Large nonspecific products (>1kb) were detected in negative control lanes for CoxIII. Considering that the size of this band was much larger than that of our desired product, we conclude that it is nonspecific and therefore it neither represents the asRNA nor the mRNA of our target genes. A future objective is to isolate these bands for cloning and sequencing to determine the identity of these bands and confirm our conclusion or modify our interpretation. For all tested genes, Lane 6 contained a PCR product of expected size indicating the presence of the relevant mRNA. Similarly, lane 8 provided evidence for the existence of asRNAs as indicated by the presence of a specific product in this lane for the nuclear, mitochondrial and chloroplast genes. The results confirm that asRNAs are present and expressed in K. brevis cells, and the same asRNAs present in the EST clusters were detectable via strand-specific PCR. I conclude that asRNAs are real transcripts and not artifacts of previous cloning and PCR experiments. The McLean EST library contains over one hundred clusters that have one or more asRNAs within them (Pirooznia et al., in press). Given the positive identification of asRNAs in the subset of genes we tested; three nuclear genes, two mitochondrial and 2 putative chloroplast genes, we predict that asRNAs are a global phenomenon in the Karenia brevis transcriptome.

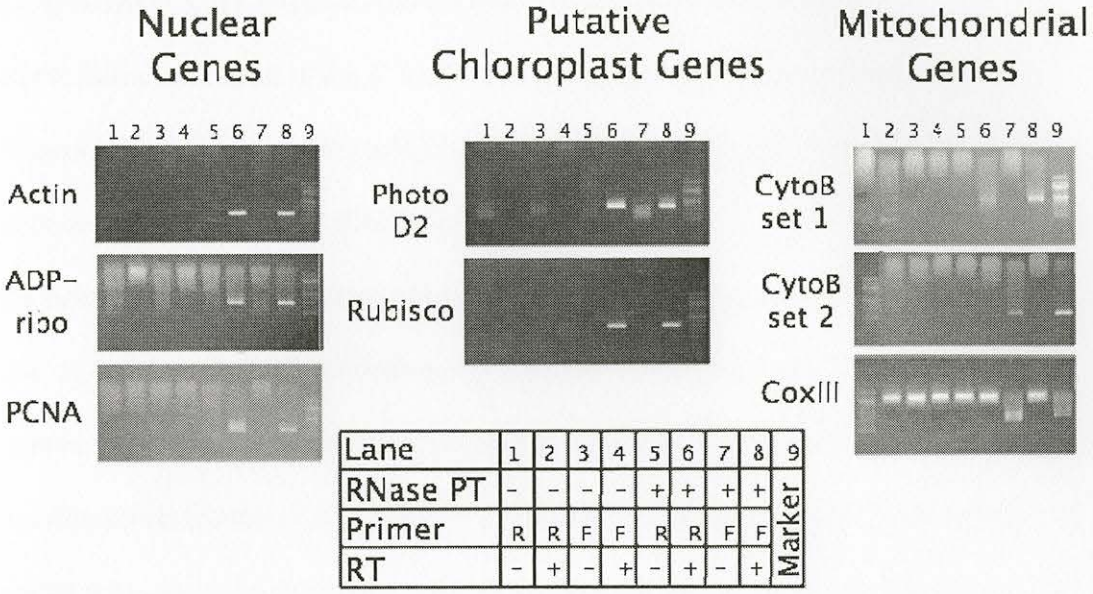


Figure 4. Strand-specific PCR Confirms Independent Existence of asRNAs and mRNAs for Nuclear, Chloroplast and Mitochondrial Genes in *Karenia brevis*. RNA was incubated with RNase to degrade all RNA in an RNase pre-treatment (RNase PT) reaction (Lanes 1-4). RNA was reverse transcribed with a strand/gene specific primer in the absence of reverse transcriptase (RT): Lanes 5+7 (negative control) or in the presence of RT (test samples): Lanes 6+8. Bands denote the presence of mRNA (Lane 6) or asRNA (Lane 8). Ladder in each gel (Lane 9) is 100bp (NEB, Ipswich, MA). CytoB set 2 and Cox III ladders precede sample lanes.

RT- and RACE-PCR

Full length sequences were generated for ADP-ribosylation factor's mRNA and asRNA sequences (Figure 5A). For photo D2, we captured the full length sequences for the asRNA strand and the 3' end of the mRNA sequence (Figure 5B). RACE-PCR was successfully used to generate the full mRNA sequence for actin and the 5' end of the asRNA (Figure 5C). Sequence data was not obtained for either the 5' end of photo D2 mRNA or the 3' end of actin asRNA although efforts to obtain those ends are continuing.

Analysis of the 5' ends of retrieved mRNAs showed the existence of the conserved Karenia brevis spliced leader sequence

UCCGUAGCCAUUUUGGCUCAAG. This 22 nucleotide fragment sits upstream of the initiation codon in the 5' UTR. Additionally, the 3' ends of both ADP-ribosylation factor and actin mRNAs indicated the presence of a stop codon and a poly-A tail. Although our RACE products for the 3' end of photo D2 did not include the poly A tail, Karenia brevis photo D2 sequences deposited in GenBank indicated that the tail is only 8 nucleotides away from the 3' end detected in our photo D2 clones. As indicated in Figure 5C, RACE products for the 5' end of actin mRNA are not displayed. However, the entire 5' end of this sequence, including the spliced leader are present in our cloned data set but are not shown in Figure 5C due to space constraints. The 5' ends of full-length asRNAs acquired with RACE-PCR indicated the absence of a spliced leader sequence but like mRNAs, asRNA transcripts contained 3' ends with poly-A tails. All asRNA transcripts lacked open reading frames, further confirming that they are indeed ncRNAs. Unique to the 5' ends of two of the three asRNA strands was sequence variability indicating, potentiality, more than one transcription start site. Sequences of clones for the 5' end of actin asRNA fell into one of two datasets; one contaning clones that started with **UGG** (after the underlined sequence at the 5' end of actin asRNA in Figure 5C) and the other contained clones with additional sequence (the underlined sequence at the 5' end of the actin asRNA). Similarly, RACE-PCR generated products for the 5' end of photo D2 asRNA started at one of two locations; either the full-length sequence as shown or after the underlined region (Figure 5B). The presence of two initiation sites for asRNAs could also indicate two separate loci coding for these transcripts, but that scenario seems less likely since the remainder of the sequence between the two clone

sets are otherwise identical. A more likely explanation is the presence of two potential transcriptional start sites at a single locus.

The length of the region of complementarity between antisense and mRNA transcripts (indicated by asterisks in Figure 5) was determined for ADP-ribosylation factor, actin and photo D2 genes. The double stranded region for ADP-ribosylation factor stretched for 685bp of near-perfect complementarity, interrupted by only 7 sporadic non-matches. For photo D2, near-perfect complementarity extended over 411bp with only 1 G-U mismatch. Given that the full 3' end for photo D2 was not acquired, it is likely that this region of complementarity is longer. Actin's region of complementarity was 884bp long with ~ 67 intermittently placed non-matches and 14 G-U mismatches. Regions of complementarity for ADP-ribo, photo D2 and actin vary in size, with ADP-ribo forming double stranded regions that span nearly the entire coding region as well as the 3' UTR of the mRNA. Both photo D2 and actin have regions of complementarity that only cover the 3' end of the mRNA but those regions stretch past the stop codon right up to the poly-A tail of their respective mRNAs. When hypothesize that asRNAs that bind mRNAs in this manner will certainly affect the rate and/or efficiency of translation of the mRNAs. Additionally, it has been posited that the presence of a spliced leader sequence on mRNAs is a signal for translation of those transcripts to proceed. Therefore, the absence of spliced leader sequences in asRNAs prevents them from being expressed or even competing for any potential translation initiation machinery that may be targeted to binding to the spliced leader sequence. Interestingly, asRNAs have poly-A tails like mRNAs. It appears that asRNAs undergo some of the post-transcriptional processing, i.e. poly-A tail addition, of mRNAs but not all, i.e. spliced leader addition. How the cells

discriminate between sense and antisense transcripts for trans-splicing will require further experimentation. The potential formation of dsRNAs between complementary asRNA and mRNA, the presence of a spliced leader sequence on mRNAs only, and the high likelihood of both features to affect translation all provide support for a post transcriptional model of gene regulation.

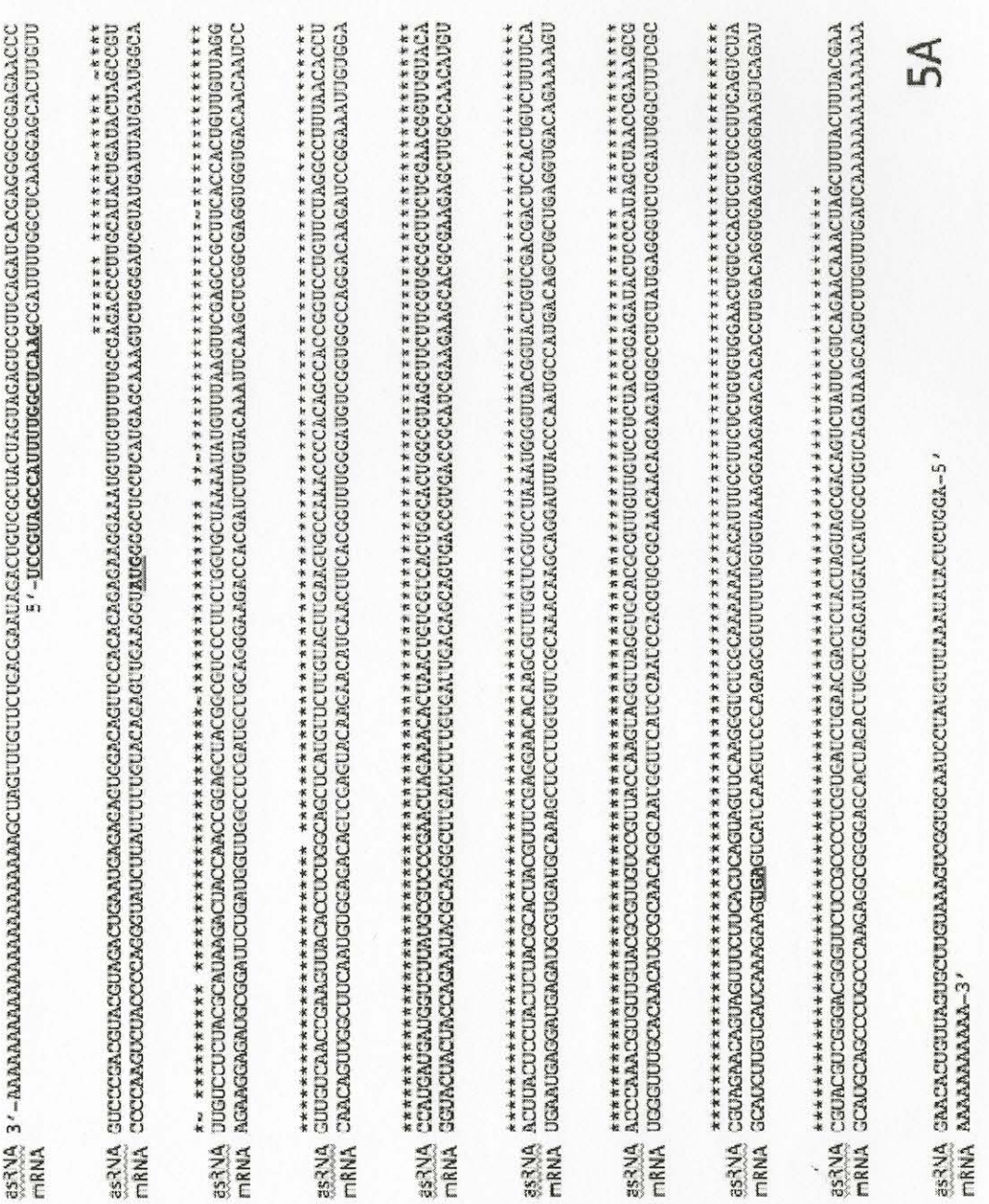


Figure 5A. RACE and RT-PCR Data Generated for mRNA and asRNA Transcripts for ADP-ribosylation factor. For full-length mRNA sequences, the 22nt splieced leader sequence at the 5' end, the initiation AUG codon, and the stop codon are underlined and represented in bold. For asRNAs, variability of transcriptional start sites is shown at the 5' end with one start site producing additional sequence (underlined and in bold) and the second site beginning just after the underlined and bolded nucleotides. Asterisks indicate regions of complementarity between mRNA and asRNA strands. Gaps in the asterisks indicate non-matches and tildes represent G-U matches.

Figure 5B. RACE and RT-PCR Data Generated for mRNA and asRNA Transcripts for Photosystem D2 gene. For full-length mRNA sequences, the 22nt splieced leader sequence at the 5' end, the initiation AUG codon, and the stop codon are underlined and represented in bold. For asRNAs, variability of transcriptional start sites is shown at the 5' end with one start site producing additional sequence (underlined and in bold) and the second site beginning just after the underlined and bolded nucleotides. Asterisks indicate regions of complementarity between mRNA and asRNA strands. Gaps in the asterisks indicate non-matches and tildes represent G-U matches.

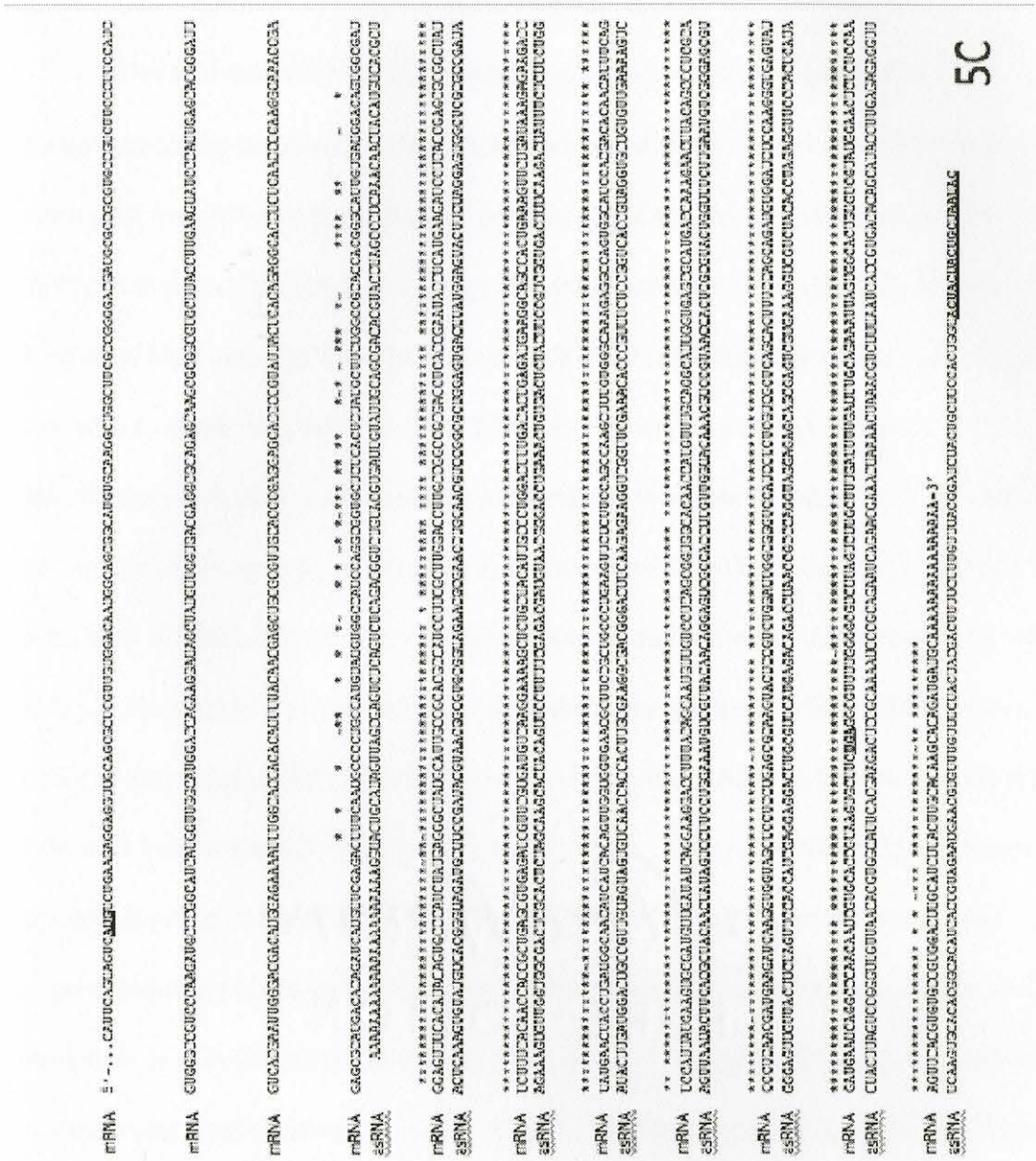


Figure 5C. RACE and RT-PCR Data Generated for mRNA and asRNA Transcripts for actin genes. For full-length mRNA sequences, the 22nt splieced leader sequence at the 5' end, the initiation AUG codon, and the stop codon are underlined and represented in bold. For asRNAs, variability of transcriptional start sites is shown at the 5' end with one start site producing additional sequence (underlined and in bold) and the second site beginning just after the underlined and bolded nucleotides. Asterisks indicate regions of complementarity between mRNA and asRNA strands. Gaps in the asterisks indicate non-matches and tildes represent G-U matches.

Immunofluorescence

The fact that asRNAs and mRNAs have such extensive regions of complementarity implies that they likely hybridize in vivo and form dsRNAs. I attempted to confirm this prediction by looking for evidence of the presence of dsRNAs inside cells. I took advantage of a commercially available monoclonal antibody, J2, that is specific for dsRNAs and has been used extensively in virology research to detect the presence of dsRNA viral genomes and replication intermediates. Figure 6 represents experiments in which cells were incubated with the secondary antibody only as a negative control (panels A-D), cells were incubated with both primary and secondary antibodies to detect the presence of dsRNAs (panels E-H), cells were incubated with both antibodies but pre-treated with RNAse III in order to degrade dsRNAs (panels I-L), and cells were incubated but pre-treated with RNase A to degrade ssRNAs (panels M-P). Panels A, E, I and M are light micrograph images showing the outline of Karenia brevis cells in the absence of filters. The control experiments confirm that any signal identified for dsRNAs was due to detection of only dsRNAs (M-P). Panels B, F, J and N represent autofluorescence caused when a signal is detected from emission produced by excitation of chlorophyll in the cells' chloroplasts. A green signal is detected from emission at a wavelength of LP 505nm. Therefore, the nucleus is visible in these panels because propidium iodide emits a signal in the same spectrum as the chlorophyll contained in the chloroplasts. A signal was detected from emission of red fluorescent molecules (not pictured) at a wavelength of LP 560nm. Longpass (LP) filters transmit wavelengths over the range of the target spectrum e.g visible spectrum (Acute Optical Technology, 2007). Panels

C, J, K and O show the fluorescence channel that detects emitted signals from fluorophores such as AMCA. The purpose of this channel is to show the blue fluorescent signal emitted by AMCA at a wavelength of BP 420-480nm. Bandpass or BP filters only transmit wavelengths within a certain range (Acute Optical Technology, 2007), therefore only the AMCA signal is detected. Under these given treatments, panels D, H, L and P are merged images of the light micrograph, autofluorescence and the AMCA signal. Presence of signal in Panel G (no RNase pre-treatment and incubated with both primary and secondary antibodies) and O (pre-treated with RNase A and incubated with primary and secondary antibodies) indicates presence of dsRNAs. As a control, some cells were incubated in the secondary antibody only with the expectation that dsRNAs would not be identified and a signal would not be detected since the secondary antibody would have nothing to bind to (the absence of any bound primary antibody). As expected, panel C shows a lack of signal. To confirm that the signal acquired represents presence of dsRNAs only, two RNase controls were performed: an RNase III control in which cells were treated with RNase III to degrade dsRNAs (Panels I-L) and an RNase A control where cells were incubated with RNase A to break down single-stranded RNAs (Panels M-P). As expected, with all dsRNAs degraded, no fluorescent signal was identified for cells pre-treated with RNase III (Panel K) while a signal was detected for those cells that were pre-treated with RNase A (Panel O). Panel L shows that the lone signal detected for the RNase III samples was due to extracellular debris and not dsRNAs inside the cell. These results support our hypothesis that dsRNAs are formed in Karenia brevis cells between the complementary regions of mRNA and asRNAs. Undeniably, the

presence of other dsRNA molecules such as transfer (tRNAs) and ribosomal RNAs (rRNAs) contribute to the fluorescent signal from detection of dsRNAs. However, rRNAs and tRNAs are localized to ribosomes. As shown by widespread detection of AMCA signal (in the nucleus and throughout the cytoplasm), our results (Panel G+O) suggest a global presence of dsRNAs and indicate the presence of dsRNAs other than tRNAs and rRNAs.

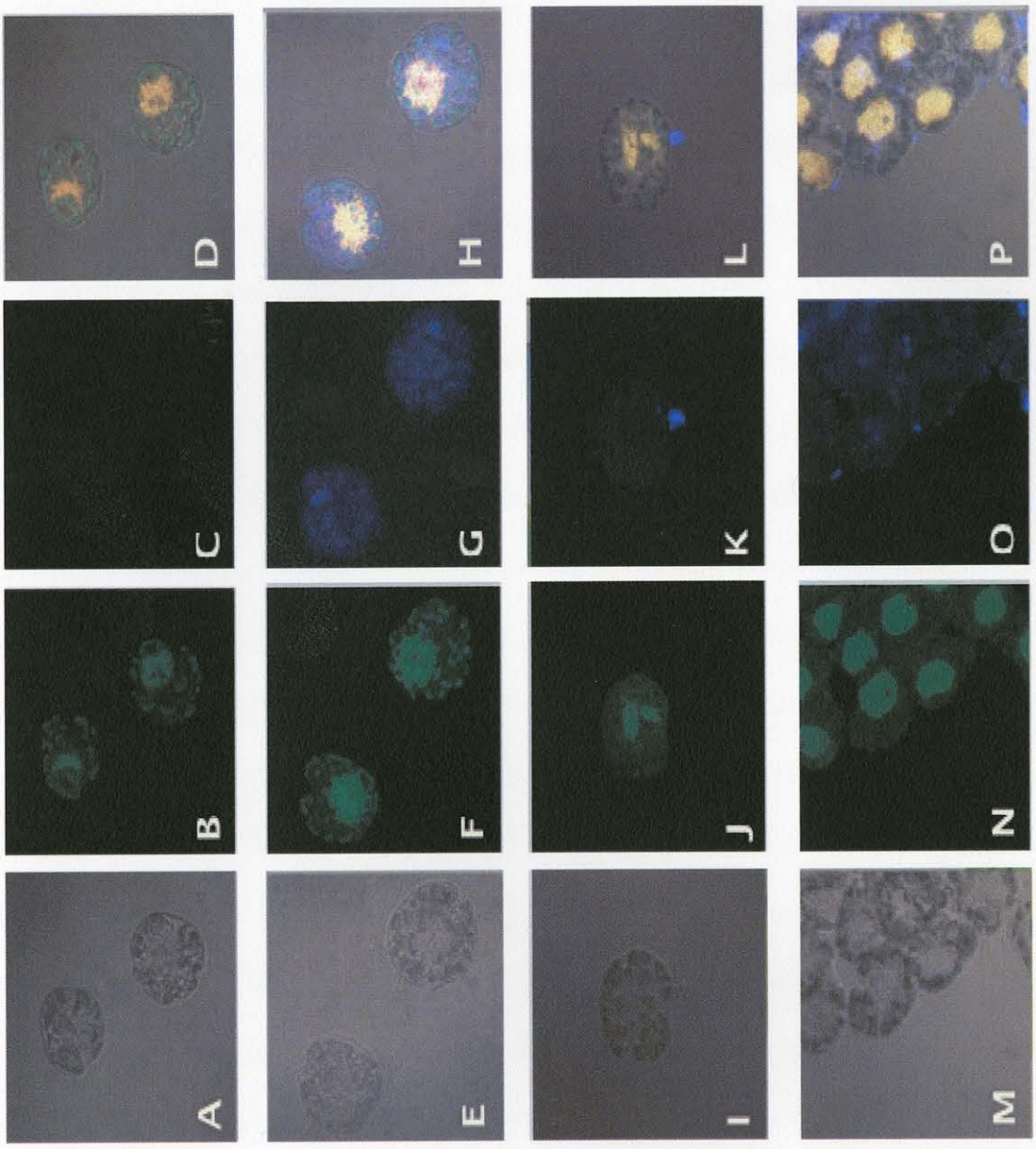


Figure 6. Immunofluorescence Images Showing Double-stranded RNA in Karenia brevis Cells. Cells were incubated with secondary antibody only (A-D); primary and secondary antibodies (E-H); Pre-treated with RNase III to degrade dsRNA (I-L) and pre-treated with RNase A to degrade ssRNA (M-P): Panels A, E, I and M indicate light micrograph images of clearly delineated K.brevis cell structures without the use of filters. Autofluorescence from chloroplasts is detected by emission of a signal in the green range at LP 505nm and nuclei are visualized at the same wavelength (B, F, J, N). Cells incubated with both primary and secondary antibody, goat anti-mouse antibody carrying an AMCA fluorophore that is detected in the blue range, BP 420-480nm (G+O) indicate presence of dsRNAs. As shown in Panel L, the blue signal detected in Panel K is due to extracellular material. Panels D, H, L and P represent merged light micrograph, autofluorescence and detected AMCA fluorophore signals.

QPCR

More often than not, conventional PCR methods yield semi-quantitative results that are based on visual quantification and although reliable quantification may be attained using these and other procedures, the protocols involved are often time and resource consuming (Bautista, Yagubi, & Roberts, n.d.; Szajek, 2007). Any efforts at quantitative analysis such as gel densitometry and hybridization assays, are only achieved after the PCR run (Hunt, 2010). In addition to the protracted process, traditional PCR assumes a reaction efficiency of 100%. This would mean that PCR products are perfectly doubled from one cycle to next until the end of the reaction and we know that this is not the case (Bar et al., 2011; BioRad, 2012; Szajek, 2007). The efficiency of a PCR describes the kinetics during the reaction where a PCR run is separated into three phases; exponential, linear and plateau. Bar et al. (2011) defines PCR kinetics as the change in DNA concentration along several consecutive cycles during the entire PCR reaction. Consequently, DNA concentration during the exponential phase, where the reaction is most accurate, shows exact doubling of PCR product. In the linear phase, the reaction starts to slow down as reaction components become limiting and amplicons start to degrade. Finally, as no more products are synthesized, the reaction stops in the plateau phase. PCR kinetics will differ from one reaction to another (even replicates) and by the end of the PCR run, reactions may have different product concentrations. In conventional PCR methods, quantifications are based on this plateau phase resulting in substandard assessments. On the other hand, in qPCR, measurements are taken when PCR amplification efficiency is exponential resulting in more accurate quantification. Moreover, the software used in

determining nucleic acid concentrations can be programmed to use mathematical models that account for any further inconsistencies (Bar et al., 2011; Pfaffl, 2001). We used the ΔΔ Ct method that is used to normalize a target gene to a reference gene.

I compared the levels of mRNA and asRNA transcripts for two genes in cells collected during the day and night (Figure 7). Considering the error bars, levels of photo D2 asRNA at night were statistically equivalent to asRNA levels during the day. Similarly, mRNA levels remained statistically stable in night versus day samples and were not that different from asRNA levels. Because levels remained unchanged for asRNA and mRNA transcripts over the circadian cycle, it is unlikely that asRNAs have any influence on the expression levels of their mRNA counterparts. Additionally, these results indicate that photo D2 does not respond to day versus night conditions, a decidedly unexpected reaction for a photosynthetic gene. We predict that under other variable conditions such as growth in different salinity or temperature, this gene may show evidence of differential expression for one or both transcripts, but we did not see those results here.

Like photo D2, levels of actin mRNA transcripts were relatively stable between day and night samples. asRNA levels in the day were comparable to mRNA levels, however, in comparison to transcript levels during the day, asRNAs for photo D2 increased dramatically, roughly three fold, at night. Research on several well studied organisms have revealed that actin levels are fairly consistent under different conditions, and in many of these studies, actin has been successfully used as a reference gene (Mori, Wang, Danenberg, Pinski, & Danenberg, 2008; Scharlaken et al., 2008) which is not inconsistent with my findings for the mRNA levels. It is

surprising that my results yielded distinctly differing expression levels for asRNA actin. Given our hypothesis that asRNAs and mRNAs bind at their complementary regions to form dsRNAs, it would seem that in this case, an excess of asRNAs at night ensures that mRNAs are not transcribed. Since it is unclear at this point why Karenia brevis cells would preferentially inhibit translation of actin mRNA during the night, this finding merits further study.

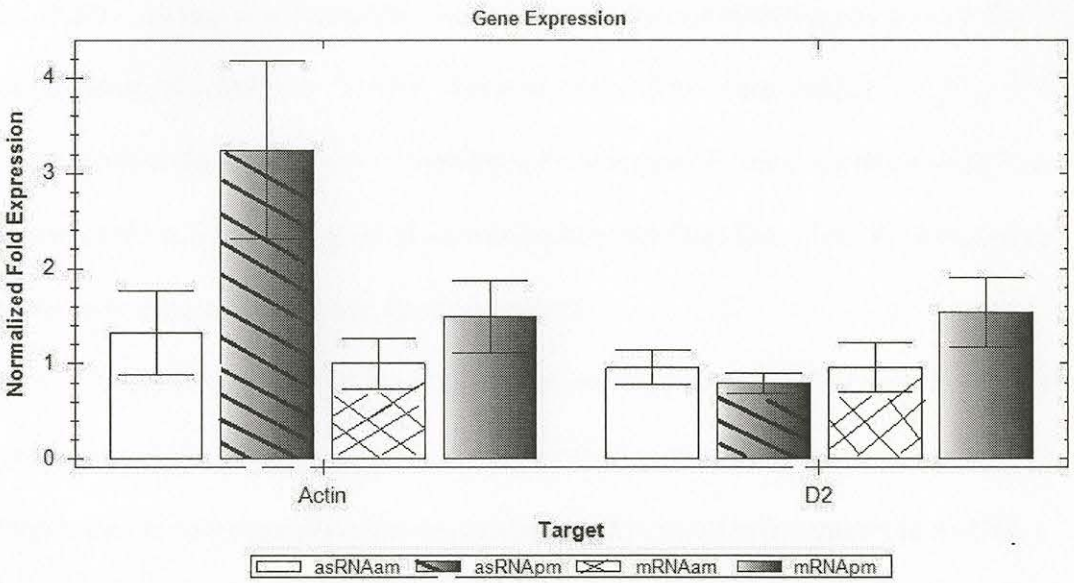


Figure 7. Differential Expression of asRNAs and mRNAs under Day vs. Night Conditions for Actin and Photosystem D2. RNA was harvested at 9AM and 9PM for day and night samples, respectively. DNase treated RNA was reverse transcribed with gene/strand specific primers and then amplified during quantitative PCR with PCNA as the reference gene. Expression levels for actin and photo D2 were normalized to PCNA. Actin asRNA day: ~1.2; Actin asRNA night: ~3; Actin mRNA day+night: ~ 1.4; D2 asRNA day+night: ~1; D2 mRNA day+ night: ~1.

CHAPTER V

SUMMARY AND CONCLUSIONS

Dinoflagellates are some of the most abundant planktonic species in the marine ecosystem, second only to diatoms. They contribute considerably to the aquatic food chain as primary producers and are also important grazers of marine bacteria and other eukaryotic plankton. A small number of dinoflagellates are also responsible for the formation of harmful algal blooms (HABs), many of which are toxic (Hackett et al., 2005). As one such example, Karenia brevis poses numerous threats to both environmental and human health. Because of the difficulties in applying molecular techniques to the study of dinoflagellate (if not impossibilities, e.g there is no way to transfect these cells), progress in understanding dinoflagellates like K. brevis at the molecular genetic level is greatly undermined.

Identification of non-coding asRNA molecules for the first time in K.brevis, provides a plausible mechanistic basis for post-transcriptional control of gene expression in this organism. The presence of asRNAs complementary to mRNA molecules was verified in vivo with strand-specific PCR. Creation of full length sequences indicating that these are in fact distinct transcripts having unique 5' and 3' ends was achieved with RT-PCR and RACE-PCR experiments. In an effort to determine a potential role for asRNAs in K. brevis, I performed quantitative PCR studies. While the results for actin are in line with our predictions for a role of asRNAs in regulating gene expression at a post-transcriptional level, it should be noted that as a non-photosynthetic gene, these results were completely unexpected for actin. I showed that actin asRNA transcripts were expressed at significantly higher

levels than mRNAs in the absence of light (night samples). During the day, asRNA levels were statistically equivalent to mRNA levels. This suggested to us, a possible role in post-transcriptional regulation of the translation machinery where actin asRNAs may affect translatability of actin mRNA. In order that more indisputable results may be achieved, clearly more candidate genes need to be assessed and other conditions (temperature and salinity) considered. Finally, we were able to show the presence of double stranded RNA (dsRNA) in Karenia cells. Having predicted that asRNAs complex with their mRNA counterparts to form dsRNA molecules, our immunofluorescence studies provide the first evidence for visualization and distribution of dsRNA in K. brevis. To support our hypothesis. current experiments in the McLean lab have been successful in precipitating dsRNAs.

Based on the hypothesis that asRNAs are involved in gene regulation by forming dsRNAs with their complementary mRNAs, we have produced a putative model for protein expression. Although photo D2 did not yield the expected results, we expect that other photosynthetic genes would be responsive to day versus night conditions with clear difference in expression levels. For a typical photosynthetic gene, we would expect to see lower levels of expression for asRNAs during the day in comparison to the levels of expressed mRNAs at the same time condition. At night, asRNAs would be present at higher expression levels than mRNAs. Conversely, as illustrated in Figure 8, asRNAs could be expressed in constant amounts both during the day and at night while mRNA levels would fluctuate to equal asRNA levels at night and increase during the day for a photosynthetic gene. Levels of expression for asRNA and mRNAs would be constant for a non-photosynthetic gene regardless of

the time of day. Figure 8 shows that a difference in transcript levels would translate into differences in protein expression during the day versus the night. During the day, when mRNA levels are higher than asRNA levels, actin protein would be made. During the night however, equivalent levels of asRNAs and mRNAs would prevent protein synthesis. We postulate that antisense RNAs would bind mRNAs to form dsRNAs that would impede the translation machinery. Future work on dsRNA in Karenia brevis will explore this model for transcriptional regulation and may provide evidence in support of our model for protein expression in K. brevis genes under different conditions.

Several biological processes in K. brevis are under circadian control and many of these cellular processes, including vertical migration and the vegetative cell cycle, are essential to the dinoflagellate's persistence during a HAB event (Van Dolah et al., 2008). As a follow up to our findings, future experiments will need to be conducted over time courses. At each time point, cells will be harvested to extract RNA for quantification of relevant transcripts as well as to extract proteins for immunoblot analysis with the ultimate goal of correlating transcript and protein expression levels. Additionally, or alternatively, cultures may be treated differentially, e.g. with distinct light or temperature regimes, to establish whether these changes can cause differential expression of the mRNA/asRNA pairs and their corresponding proteins. In addition, harvesting of dsRNA populations from cells under variable conditions and sequencing the members of the populations will provide additional insight into what genes are found in dsRNA complexes and how

that might relate to genetic and physiological control of those cells. Another logical direction for future experiments would be determining how Karenia brevis cells discriminate between asRNA and mRNA transcripts. Based on results from our RACE-PCR experiments, it appears that asRNAs do not have the otherwise ubiquitous 22nt spliced leader sequence. Research perfomed by other labs suggests that addition of a spliced leader is essential for translation and may provide evidence for post-transcriptional gene regulation (Lidie & Van Dolah, 2007a; Lin, 2011; Zhang et al., 2007). Therefore, it would be beneficial to identify genes that follow this predicted model.

As we have seen, Karenia brevis research already has a wide-ranging scientific reach from phylogenetics to cancer research to pharmacology (Lidie et al., 2005; Radwan & Ramsdell, 2007; Saldarriaga et al., 2003; Yoon et al., 2005). Basic questions regarding bloom dynamics and toxin production are still unanswered, however. Additional research into the molecular genetics of *K. brevis* is warranted as a potentially fruitful avenue to address these questions and more. Certainly, as investigations in the K.brevis transcriptome continue, identification of novel post-transcriptional and translational models of regulation of novel translational controls in this dinoflagellate could provide a model for similar exploration in related organisms. These control mechanisms may even produce targets for biotechnological intervention that may help alleviate the problems of K. brevis HABs.

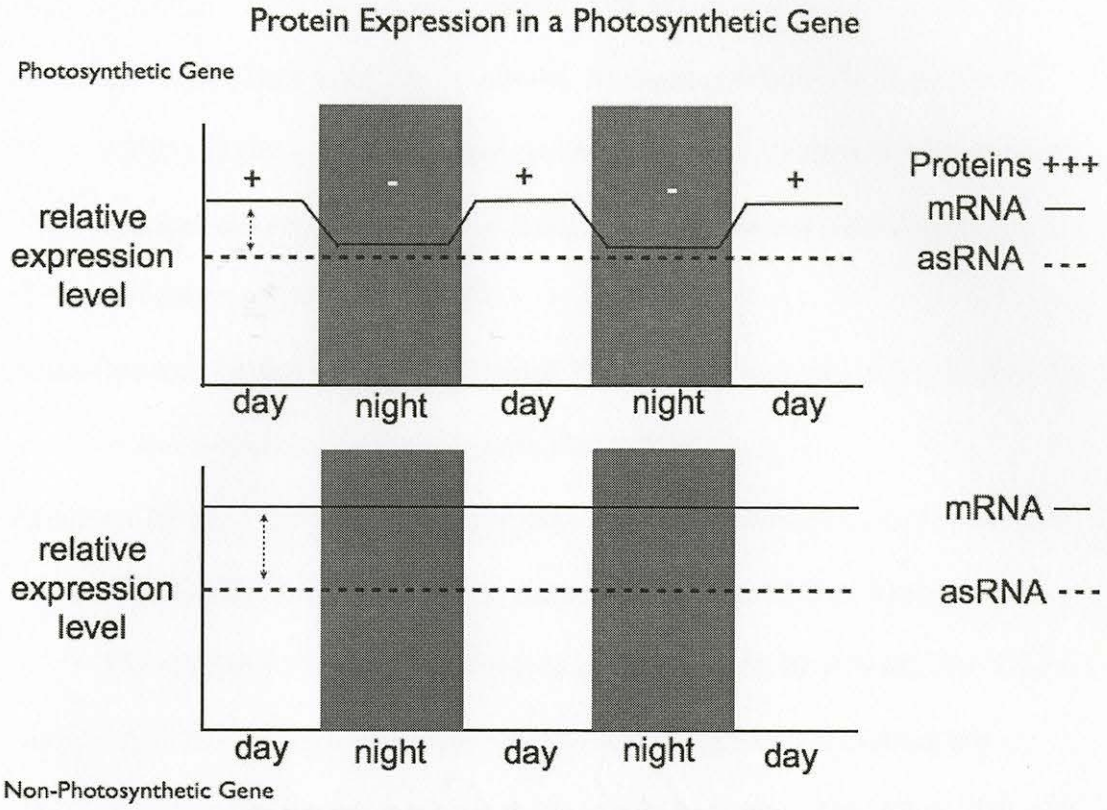


Figure 8. Predicted Model for Protein Expression under Day vs. Night Conditions for a Photosynthetic Gene and a Non-photosynthetic Gene. Day: high mRNA levels for a photosynthetic gene translate into high protein expression levels (+); Night: Decreased mRNA levels for the photosynthetic gene signify low/ no protein expression (-); Day+Night: Constant mRNA and asRNA levels translate into constant protein expression levels in a non-photosynthetic gene.

REFERENCES

Abraham, W. M., Bourdelais, A. J., Ahmed, A., Serebriakov, I., & Baden, D. G. (2005). Effects of Inhaled Brevetoxins in Allergic Airways: Toxin-Allergen Interactions and Pharmacologic Intervention. *Environmental Health Perspectives, 113*(5), 632-637.

Acute Optical Technology. (2007). Optical Filters. Retrieved July 8, 2012 from http://www.accuteoptical.com/UV_pass_filters.html

Anderson, D. M. (1989). Toxic algal blooms and red tides: A global perspective. In T. Okaichi, D. M. Anderson, & T. Nemoto (Eds.), Red Tides: Biology, Environmental Science and Toxicology (pp. 11-16). New York, NY: Elsevier.

Archibald, J. M. (2005). Jumping genes and shrinking genomes: Probing the evolution of eukaryotic photosynthesis with genomics. *Life, 57*(8), 539-547.

Bachvaroff, T. R., & Place, A. R. (2008). From Stop to Start: Tandem gene arrangement, copy number and trans-splicing sites in the dinoflagellate Amphidinium carterae. *PLoS ONE, 3*(8), 1-14.

Baden, D. G., Rein, K. S., Gawley, R. E., Jeqlitsch, G., & Adams, D. J. (1994). Is the α-ring lactone of brevetoxin PbTx-3 required for sodium channel orphan receptor binding and activity? *Natural toxins, 2*(4), 212-221.

Baden, D. G., Bourdelais, A.J., Jacocks, H., Michelliza, S., & Naar, J. (2005). Natural and Derivative Neurotoxins: Historical Background, Multiplicity and Effects. *Environmental Health Perspectives, 113*(5): 621-625.

Bar, T., Kubista, M., & Tichopad, A. (2011). Validation of kinetics similarity in qPCR. *Nucleic Acid Research, 40*(4), 1395-1406.

Barber, J. (2002). Photosystem II: A multisubunit membrane protein that oxdises water. *Current Opinion in Structural Biology, 12*, 523-530

Baserga, R. (1991). Growth regulation of the PCNA gene. *Journal of Cell Science, 98,* 433-436.

Bautista, D., Yagubi, A., & Roberts, P. (n.d.). Application Note 8, Molecular Weight Markers: DNA Quantification by Gel Densitometry with Norgen DNA Ladders. Retrieved May 6, 2012 from http://www.norgenbiotek.com/display-resource

Biochemistry. (2006). Differences in mechanism of transcription between bacteria and eukaryotes (PDF document). Retrieved May 6, 2012 from Lecture Notes Online: http://www.smbs.buffalo.edu/bch/Courses/bms503/sinha_s_class2.pdf

BioRad. (2012). Quantitative PCR analysis: Relative quantification. Retrieved July 8, 2012 from http://www3.bio-rad.com/B2B/vanity/gexp/content.do

Brand, L. E., & Compton, A. (2006). Long-term increase in Karenia brevis abundance along the Southwest Florida Coast. *Harmful Algae, 6*(2), 232-252.

Brown, A. F., Dortch, Q., Van Dolah, F. M., Leighfield, T. A., Morrison, W., Thessen, A. E.,… Pennock, J. R. (2005). Effect of salinity on the distribution, growth, and toxicity of Karenia spp. *Harmful Algae, 5*(2), 199-212. doi: 10.1016/j.hal.2005.07.004.

Brunelle, S. A., & Van Dolah, F. M. (2011). Post-transcriptional Regulation of S-Phase Genes in the Dinoflagellate, Karenia brevis. *Journal of Eukaryotic Microbiology, 58*(4), 373-382.

Casper, E. T., Paul, J. H., Smith, M. C., & Gray, M. (2004). Detection and quantification of the red tide dinoflagellate Karenia brevis by real-time nucleic acid sequence-based amplification. *Applied and Environmental Microbiology, 70*(8), 4727-4732.

Central Florida News 13. (2009). Red Tide, Fish Kill Affecting Florida's Seafood Market. Retrieved August 7, 2011 from http://www.cfnews13.com/content/news/cfnews13/news/article.html

Chalfie, M., Horvitz, H. R., & Sulston, J. E. (1981). Mutations that lead to reiterations in the cell lineages of C. elegans. *Cell 24*, 59–69

Chang, D. E., Smalley, D. J., & Conway, T. (2002). Gene Expression Profiling of Escherichia coli Transitions: An Expanded Stringent Response Model. *Molecular Microbiology, 45*(2), 289-306.

Cooper, G. M. (2000). The cell: A molecular approach. Sunderland, MA: Sinauer Associates.

Derby, M. L., Galliano, M., Krzanowski, J. J., & Martin, D. F. (2002). Studies of the effect of □-APONIN from Nannochloris sp.on the Florida red tide organism Karenia brevis. *Toxicon, 41*, 245–249.

Dhingra, A., Portis, A. R., & Daniel, H. (2004). Enhanced translation of a chloroplast-expressed RbcS gene restores small subunit levels and photosynthesis in nuclear RbcS antisense plants. *Proceedings of the National Academy of Sciences, 101*(16), 6315-6320.

Dominguez, R., & Holmes, K. C. (2011). Actin structure and function. *Annual Review of Biophysics, 40*, 169-186

Donaldson, J., & Segev, N. (2009). Regulation and coordination of intracellular trafficking. *Landes Bioscience*, 329-338

Eddy, S. R. (2001). Non-coding RNA genes and the Modern RNA world. *Nature Reviews: Genetics, 2*, 919-929.

Fish and Wildlife Research Institute. (2009). Effects of Florida Red Tide on Marine Animals. Retrieved July 8, 2012 from http://research.myfwc.com/features/view_article.asp

Fleming, L. E., Kirkpatrick, B., Backer, L. C., Bean, J. A., Wanner, A., Dalpra, D.,... Baden, D. G. (2005). Initial Evaluation of the Effects of Aerosolized Florida red tide Toxins (Brevetoxins) in Persons with Asthma. *Environmental Health Perspectives, 113*(5), 650-657.

Fleming, L. E., Kirkpatrick, B., Backer, L. C., Bean, J. A., Wanner, A., Reich, A.,... Baden, J. G. (2007). Aerosolized red-tide toxins (brevetoxins) and asthma. *CHEST, 131*(1), 187-194.

Flewelling, L. J., Naar, J. P., Abbott, J. P., Baden, D. G., Barros, N. B., Bossart, G. D., & Botterin, M. D. (2005). Red tides and marine mammal mortalities: Unexpected brevetoxin vectors may account for deaths long after or remote from an algal bloom. *Nature, 435*(7043), 755-756. doi:10.1038/nature435755a.

Frohman, M. A., Dush, M. K., & Martin, G. R. (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences, 85*, 8998–9002.

Funaki, K., Katsumoto, T., & Iino, A. (1995). Immunocytochemical localization of actin in the nucleolus of rat oocytes. *Biology of the Cell, 84*, 139-146

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B.,... Brown, P. O. (2000). Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes. *Molecular Biology of the Cell, 11*, 4241-4257.

Gilbert, P. M., Burkholder, J. M., Kana, T. M., Alexander, J., Skelton, H., & Shilling, C. (2009). Grazing by Karenia brevis on Synechococcus enhances its growth rate and may help to sustain blooms. *Aquatic Microbial Ecology, 55*, 17-30.

Gottesman, S. (2005). Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends in Genetics, 21*(7), 399-404.

Griffiths, A. J., Miller, J. H., Suzuki, D. T., Lewontin, R. C., & Gelbart, W. M. (1999). *Regulation of Gene Transcription. An Introduction to Genetic Analysis*. New York, NY: W.H. Freeman and Company.

Guiry, M. D., & Guiry, G. M. 2012. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. Retrieved April 9, 2012 from http://www.algaebase.org.

Gupta, J. D., Gu, H., Chernokalskaya, E., Gao, X., & Schoenberg, D.R. (1998). Identification of two cis-acting elements that independently regulate the length of poly (A) on Xenopus albumin pre-mRNA. *RNA, 4*(7), 766-776.

Hackett, J. D., Anderson, D. M., Erdner, D. L., & Bhattacharya, D. (2004). Dinoflagellates: A remarkable evolutionary experiment. *American Journal of Botany, 91*(10), 1523-1534.

Hackett, J. D., Scheetz, T. E., Yoon, H. S., Soares, M. B., Bonaldo, M. F., Casavant, T. L., & Bhattacharya, D. (2005). Insights into a dinoflagellate genome through expressed sequence tag analysis. *BMC Genomics, 6*(80). doi: 10.1186/1471-2164-80.

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nature Structural & Molecular Biology, 11*, 394-403

Hallegraeff, G. M., Valentine, J. P., Marshall, J., & Bolch, C. J. (2004). Temperature tolerances of toxic dinoflagellate cysts: application to the treatment of ships' ballast water. *Aquatic Ecology, 31*(1), 47-52.

Harata, M., Karwan, A., & Wintersberger, U. (1994). An essential gene of Saccharomyces cerevisiae coding for an actin-related protein. *Proceedings of the National Academy of Sciences, 91*, 8258-8262.

Heimann, K. (2012, March 15). *Gymnodinium* and Related Dinoflagellates. *eLS*. doi: 10.1002/9780470015902.

Hoagland, P., Jin, D., Polansky, L. Y., Kirkpatrick, B., Kirkpatrick, G., Fleming, L.,… Backer, L. C. (2009). The costs of respiratory illnesses arising from Florida Gulf Coast Karenia brevis blooms. *Environmental Health Perspectives, 117*(8), 1239-1243.

Howe, C. J., Nisbet, R. E. R., & Barbrook, A. C. (2008). The remarkable chloroplast genome of dinoflagellates. *Journal of Experimental Botany, 59*(5), 1035-1045.

Hunt, M. (2010). Microbiology and Immunology Online: Real Time PCR. Retrieved July 8, 2012 from http://pathmicro.med.sc.edu/pcr/realtime-home.htm

Jacquier, A. (2009). The compex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nature Reviews Genetic, 10*, 833-844.

Jansen, M., De Moor, C. H., Sussenbach, J. S., & Van Den Brande, J. L. (1995). Translational Control of Gene Expression. *Pediatric Research, 37*(6), 681-686.

Jeong, H. J., Park, J. Y., Nho, J. H., Park, M. O., Ha, J. H., Seong, K. A.,... Yih, W. H. (2005). Feeding by red-tide dinoflagellates on the cyanobacterium Synechococcus. *Aquatic Microbial Ecology, 41*, 131-143.

Jeong, H. J., Yoo, Y. D., Kim, J. S., Seong, K. A., Kang, N. S., & Kim, T. H. (2010). Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates in marine planktonic food webs. *Ocean Science Journal, 45*(2), 65-91.

Jonsson, Z.O., & Hubscher, U. (2005). Proliferating cell nuclear antigen: More than a clamp for DNA polymerases. *BioEssays, 19*(11), 967-975.

Komenda, J., Reisinger, V., Muller, B. C., Dobakova, M., Granvogl, B., & Eichacker, L. A. Accumulation of the D2 protein is a key regulatory step for assembly of the photosystem II reaction center complex in Synechocystis PCC 6803. *Journal of Biological Chemistry, 279*(47), 48620-48629.

Lee, D., Mittag, M., Sczekang, S., Morsell, D., & Hastings, J. W. (1992). Molecular Cloning and Genomic Organization of a Gene for Luciferinbinding Protein from the Dinoflagellate Gonyaulax polyedru. *Journal of Biological Chemistry, 26*(12), 6842-8850.

Lee, M. H., Min, M. K., Lee, Y. J., Jin, J. B., Shin, D. H., Kim, D. H., Lee, K., & Hwang, I. (2002). ADP-Ribosylation factor 1 of Arabidopsis plays a critical role in intracellular trafficking and maintenance of endoplasmic reticulum morphology in arabidopsis. *Plant Physiology, 129*, 1507-1520.

Lidie, K. B., Ryan, J. C., Barbier, M, & Van Dolah, F. M. (2005). Gene Expression in Florida Red Tide Dinoflagellate Karenia brevis: Analysis of an Expressed Sequence Tag Library and Development of DNA Microarray. *Marine Biotechnology, 7*(5), 481-493.

Lidie, K. B., & Van Dolah, F. M. (2007a). Spliced leader RNA-mediated trans-splicing in a dinoflagellate, Karenia brevis. *The Journal of Eukaryotic Microbiology, 54*(5), 427-435.

Lidie, K. B., & Van Dolah, F. M. (2007b). Characterization and regulation of gene expression networks in response to acute stress in Karenia brevis: *Fourth Symposium on Harmful Algae in the U.S.* Woods Hole, MA.

Lin, S. (2011). Genomic understanding of dinoflagellates. *Research in Microbiology, 162*, 551-569.

Lindstrom, A. (2007). Unraveling the mystery of Karenia brevis. *Florida Wildlife, 67*. Retrieved November 9, 2011 from http://floridawildlifemagazine.com/pdf/mar07-karenia.pdf

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, S. L., & Darnell, J. (2004). Transport of ions and small molecules across cell membranes. In Molecular Cell Biology (pp. 276-279). New York, NY: W.H. Freeman and Company.

Maga, G., & Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *Journal of Cell Science, 116*, 3051-3060.

Martin, D. F., & Martin, B. B. (1976). Red Tide, Red Terror: Effects of Red Tide and Related Toxins. *Journal of Chemical Education, 53*(10), 614-617.

Matthews, S. & Brandt, K. (2004). Africa Invaded: The growing danger of invasive alien species. Cape Town, RSA: Global Invasive Species Programme.

McClean, P., & Johnson, C. (n.d.) Photosystem II. Virtual Cell Animation Collection: Molecular and Cellular Biology Learning Center. Retrieved from http://vcell.ndsu.edu/animations/photosystemII/index.htm

McKay, L., Kamykowski, D., Milligan, E., Schaeffer, B., & Sinclair, G. (2006). Comparison of swimming speed and photophysiological responses to different external conditions among three Karenia brevis strains. *Harmful Algae, 5*, 623–636.

McPherson, B. F. (2003). Southern Florida NAWQA program. Retrieved November 8, 2011 from http://fl.water.usgs.gov/Sofl.fact.html.

Melis, A. (1999). Photosystem-II damage and repair cycle in chloroplasts: what modulates the rate of photodamage in vivo. *Cell, 4*(4), 130-135.

Miles, B. W. (2003). Photosystems I and II. Texas A & M University, Department of Biochemistry and Biophysics: Lecture notes. Retrieved from http://www.tamu.edu/faculty/bmiles/

Miller-Morey, J. S., & Van Dolah, F. M. (2004). Differential Responses of Stress Proteins, Antioxidant Enzymes and Photosynthetic Efficiency to Physiological

Stresses in the Red Tide Dinoflagellate, Karenia brevis. *Comparative Biochemistry and Physiology*, Part C 1*38*, 493-505.

Mittag, M., Eckerskorn, C., Strupat, K., & Hastings, J. W. (1997). Differential translational initiation of lbp mRNA is caused by a 5'upsteam open reading frame. *Federation of European Biochemical Societies Letters, 18849*(411), 245-250.

Mittag, M., Li, L., & Hastings, J. W. (1998). The mRNA Level of the Circadian Regulated Gonyaulax Luciferase Remains Constant over the Cycle. *Chronobiology International, 15*(1), 93-98.

Miyata, T., Oyama, T., Mayanagi, K., Ishino, S., Ishino, Y., & Morikawa, K. (2004). The Clamp- loading complex for processive DNA replication. *Nature Structural and Molecular Biology, 11*, 632-636.

Morey, J. S., Monroe, E. A., Kinney, A. L., Beal, M., Johnson, J. G., Hitchcock, G. L., & Van Dolah, F. M. (2011). Transcriptome response of the red tide dinoflagellate, Karenia brevis, to nitrogen and phosphorus depletion and addition. *BMC Genomics, 12*, 346- 1-18.

Mori, R., Wang, W., Danenberg, K. D., Pinski, J. K., & Danenberg, P. V. (2008). Both β-actin and GAPDH are useful reference genes for normalization of quantitative RT-PCR in human FFPE tissue samples of prostate cancer. *The Prostate, 68*(14), 1555-1560.

Morse, D., Milos, P. M., Roux, E., & Hastings, J. W. (1989). Circadian regulation of bioluminescence in Gonyaulax involvestranslational control. *Proceedings of the National Academy of Sciences, 86*, 172-176.

Morse, D. S, Fritz, L., & Hastings, J. W. (1990). What is the clock? Translational regulation of circadian bioluminescence. *Trends in Biochemical Sciences, 15* (7), 262-265.

Morse, D., Salois, P., Markovic, P., & Hastings, J. W. (1995). A nuclear-encoded form II RuBisCO in dinoflagellates. *Science, 268*(5217), 1622-1624.

National Centers for Coastal Ocean Science. (2009). Economic impacts of Harmful Algal Blooms (HABs). Retrieved from http://www.cop.noaa.gov.stressors.extremeevents/hab/current/HAB_Econ.html

National Centers for Coastal Ocean Science. (2006). State of the Research on Red Tide in the Gulf of Mexico forum seeks public input. Retrieved April 12, 2012 from http://www.cop.noaa.gov/stressors/extremeevents/hab/features/red_tide_mtg_0606.html

National Institute of Environmental Health Sciences. (1996). Analysis of biotoxins (red tide) in manatee tissues (2nd ed.). Miami, FL: Author.

Neymotin, S. B. (2011). RNA Extraction. New York University Gresham Lab Protocols. Retrieved 2012 on July 8, from https://wikis.nyu.edu/display/greshamlab/RNA+extraction

Nickelsen, J., Fleischmann, M., Boudreau, E., Rahire, M., & Rochaix, J. (1999). Identification of cis-acting RNA leader elements required for choloroplast psbD gene expression in Chlamydomonas. *The Plant Cell, 11*, 957-970.

Nosenko, T., & Bhattacharya, D. (2007). Horizontal gene transfer in chromalveolates. *BMC Evolutionary Biology, 7*(173), 1471-2148

Nosenko, T., Lidie, K. L., Van Dolah, F. M., Lindquist, E., Cheng, J., & Bhattacharya, D. (2006). Chimeric plastid proteome in the Florida "red tide" dinoflagellate Karenia brevis. *Molecular Biology and Evolution., 23*(11), 2026-2038.

Okamoto, O. K., & Hastings, J. W. (2003). Novel dinoflagellate clock-related genes identified through microarray analysis. *Journal of Phycology, 39*(3), 519-526.

Palmer, J. D. (1996). Rubisco surprises in dinoflagellates. *The Plant Cell, 8*, 343-345

Perret, E., Davoust, J., Albert, M., Besseau, L., & Soyer-Gobillard, M. (1993). Microtubule organization during the cell cycle of the primitive eukaryotes dinoflagellate Crypthecodinium cohnii. *Journal of Cell Science, 104*, 639-651.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research, 29*(9), 2002-2003.

Pierce, R. H., Henry, M. S., Blum, P. C., Lyons, J., Cheng, Y. S., Yazzie, D., & Zhou, Y. (2003). Brevetoxin Concentrations in Marine Aerosol: Human Level Exposures during a Karenia brevis Harmful Algal Bloom. *Bulletin of Environmental Contamination and Toxicology, 70*(1), 161–165. doi:10.1007/s00128-002-0170-y.

Pirooznia, M., Gibson, R., Zhang, C., Deng, Y., Namataka, H., & McLean, T. I. (in press). Identification of Antisense Transcripts in a Multi-strain EST Library for the Harmful Algal Bloom-forming Dinoflagellate, *Karenia brevis*. *Journal of Phycology*.

Plath, A., Peters, F., & Einspanier, R. (1995).Detecting mRNAs by Ribonuclease Protection Assay using denaturing horizontal polyacrylamide gel electrophoresis: A radioactive and nonradioactive approach. *Electrophoresis, 17*(3), 471-472.

Prevost, K., Salvail, H., Desnoyers, G., Jacques, E., & Masse, E. (2007). The small RNA RyhB activates the translation of shiA mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. *Molecular microbiology, 64*(5), 1260-1273.

Rabalais, N. N. (2004). Hypoxia in the Gulf of Mexico. In Environmental Analysis of the Gulf of Mexico (Management of the Coastal Zone of the Gulf). Retrieved from http://www/harteresearchinstitute.org/ebook/ch26-hypoxia.pdf

Radwan, F. Y., & Ramsdell, J. S. (2007). Brevetoxin form covalent DNA adducts in rat lung following intratracheal exposure. *Environmental Health Perspectives, 116*(7), 930-936.

Raskin, D. M., Judson, N., & Mekalanos, J. J. (2007). Regulation of the stringent response is the essential function of the conserved bacterial G protein CgtA in Vibrio cholerae. *Proceedings of the National Academy of Sciences, 104*(11), 4636-4641.

Reisler, E., & Egelman, E. H. (2007). Actin structure and function: what we still do not understand. *The Journal of Biological Chemistry, 282*(50), 36133-36137.

Saldarriaga, J. F., McEwan, M. L., Fast, N. M., Taylor, F. J. R., & Keeling, P. J. (2003). Multiple protein phylogenies show that Oxyrrhis marina and Perkinsus marinus are early branches of the dinoflagellate lineage.

International journal of Systematic and Evolutionary Microbiology, 53, 355-365.

Santiago-Vazquez, L. Z., Ranzer, L. K., & Kerr, R. G. (2006). Comparison of two total RNA extraction protocols using the marine gorgonian coral Pseudopterogorgia elisabethae and its symbiont Symbiodinium sp. Electronic *Journal of Biotechnology, 9*, 598-603.

Scharlaken, B., de Graaf, D. C., Goossens, K., Brunain, M., Peelman, L. J., & Jacobs, F. J. (2008). Reference gene selection for insect expression studies using quantitative real-time PCR: The honeybee, Apis mellifera, head after a bacterial challenge. *Journal of Insect Science, 8*(33), available online: insectscience.org/8.33

Seong, K. A., Jeong, H. J., Kim, S., Kim, G. H., & Kang, J. H. (2006). Bacterivory by co-occuring red-tide algae, heterotrophic nanoflagellates, and ciliates. *Marine Ecology Progress Series, 322*, 85-97.

Smith, C. J., & Osborn, A. M. (2008). Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbial Ecology, 67*, 6-20.

Soyer-Gobillard, M., Ausseil, J., & Geraud, M. (1996). Nuclear and Cytoplasmic actin in dinoflagellates, *Biology of the Cell, 87*, 17-35.

Szajek, A. W. (2007). Nucleic Acid Based Techniques- Approaches for Detecting Trace Nucleic Acids (Residual DNA Testing). *Pharmacopeial Forum, 33*(5), 1025.

Tsurimoto, T. (1999). PCNA Binding Proteins. *Frontiers in Bioscience, 4*, 849-859.

U.S Environmental Protection Agency (2008). National Coastal Condition Report III: Gulf Coast Coastal Condition. Retrieved from http://www.epa.gov/nccr

Van Dolah, F. M., Lidie, K. B., Monroe, E. A., Bhattacharya, D., Campbell, L., Doucette, G. J.,… Kamykowski, D. (2008). The Florida red ride dinoflagellate Karenia brevis: New Insights into cellular and molecular processes underlying bloom dynamics. *Harmful Algae, 8*(4), 562-572. doi: 10.1016/j.hal.2008.11.004.

Van Guilder, H. D., Vrana, K. E., & Freeman., W. M. (2008). Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques, 44*(5), 619-626.

Vogel, J., Bartels, V., Tang, T. H., Churakov, G., Slagter-Jager, J. G., Huttenhofer, A., & Wagner, E. G. H. (2003). RNomics in Eshercichia coli detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Research, 32*(22), 6435-6443.

Walker, L. M. (1982). Evidence for a Sexual Cycle in the Florida Red Tide Dinoflagellate, Ptychodiscus brevis (Gymnodinium breve). *Transactions of the American Microscopical Society, 101*(3), 287-293.

Woodcock, A. H. (1948). Note concerning human respiratory irritation associated with high concentrations of plankton and mass mortalities of marine organisms. *Journal of Marine Research, 7*, 56–62.

Yang, H., & Weisberg, R. H. (1999). Response of the West Florida Shelf circulation to climatological wind stress forcing. *Journal of Geophysical Research, 104*, 5301-5320.

Yoon, H. S., Hackett, J. D., & Bhattacharya, D. (2002). A single origin of the peridinin- and fucoxanthin- containing plastids in dinoflagellates through tertiary endosymbiosis. *Proceedings of the National Academy of Sciences, 99*(18), 11724-11729.

Yoon, H. S., Hackett, J. D., Van Dolah, F. M., Nosenko, T., Lidie, K. L., & Battacharya, D. (2005). Tertiary Endosymbiosis Driven Genome Evolution in Dinoflagellate Algae. *Molecular Biology and Evolution, 22*(5), 1299-1308.

Zhao, L., Mi, T., Zhen, Y., Li, M., He, S., Sun, J., & Yu, Z. (2009). Cloning of proliferating cell nuclear antigen gene from the dinoflagellate Prorocentrum donghaiense and monitoring its expression profiles by real-time RT-PCR. *Hydrobiologia, 627*, 19-30.

Zhang, H., Hou, Y., Miranda, L., Campbell, D. A., Sturm, N. R., Gaasterland, T., & Lin, S. (2007). Spliced leader RNA trans-splicing in dinoflagellates. *Proceedings of the National Academy of Sciences, 104*(11), 4618-4623.

Zhang, Z., Cavalier-Smith, T., & Green, B. R. (2002). Evolution of dinoflagellate unigenic minicircles and the partially concerted divergence of their putative replicon origins. *Molecular Biology and Evolution, 19*(4), 489-500.

Zheng, B., Han, M., Bernier, M., & Wen, J. (2009). Nuclear actin and actin-binding proteins in the regulation of transcription and gene expression. *The FEBS Journal, 276*(10), 2669-2685.